

4121-129

U.S. APPLICATION NO. (If known, see 37 CFR 1.5)

09/914549

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		
INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/DE00/00583	28 February 2000	26 February 1999
TITLE OF INVENTION		
PROTEIN (TP) THAT IS INVOLVED IN THE DEVELOPMENT OF THE NERVOUS SYSTEM		
APPLICANT(S) FOR DO/EO/US		
Annemarie Poustka and Johannes Coy		
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:		
<p>1. <input checked="" type="checkbox"/> This is a <b>FIRST</b> submission of items concerning a filing under 35 U.S.C. 371.</p> <p>2. <input type="checkbox"/> This is a <b>SECOND</b> or <b>SUBSEQUENT</b> submission of items concerning a filing under 35 U.S.C. 371.</p> <p>3. <input type="checkbox"/> This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).</p> <p>4. <input checked="" type="checkbox"/> A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.</p> <p>5. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2))           <ul style="list-style-type: none"> <li>a. <input checked="" type="checkbox"/> is transmitted herewith (required only if not transmitted by the International Bureau).</li> <li>b. <input checked="" type="checkbox"/> has been transmitted by the International Bureau.</li> <li>c. <input type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US).</li> </ul> </p> <p>6. <input checked="" type="checkbox"/> A translation of the International Application into English (35 U.S.C. 371(c)(2)).</p> <p>7. <input type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))           <ul style="list-style-type: none"> <li>a. <input type="checkbox"/> are transmitted herewith (required only if not transmitted by the International Bureau).</li> <li>b. <input type="checkbox"/> have been transmitted by the International Bureau.</li> <li>c. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired.</li> <li>d. <input checked="" type="checkbox"/> have not been made and will not be made.</li> </ul> </p> <p>8. <input type="checkbox"/> A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).</p> <p>9. <input checked="" type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).*(<b>Unsigned</b>)</p> <p>10. <input type="checkbox"/> A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).</p>		
<b>Items 11. to 16. below concern other document(s) or information included:</b>		
<p>11. <input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98.</p> <p>12. <input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.</p> <p>13. <input checked="" type="checkbox"/> A <b>FIRST</b> preliminary amendment.  <input type="checkbox"/> A <b>SECOND</b> or <b>SUBSEQUENT</b> preliminary amendment.</p> <p>14. <input type="checkbox"/> A substitute specification.</p> <p>15. <input checked="" type="checkbox"/> A small entity statement.</p> <p>16. <input type="checkbox"/> Other items or information: EPO Search Report and International Preliminary Examination Report in German, Computer Readable Disk with Sequence Listing</p>		

NOTE: This application is being filed with an unsigned Oath or Declaration under the provisions of 37 CFR § 1.53 in order that applicant may secure a filing date of August 24, 2001. Upon receipt of a "Notice to File Missing Parts - Filing Date Granted," a executed Declaration and Power of Attorney will be forwarded. The undersigned agent affirmatively states that she has been duly authorized and appointed to file this application on behalf of the applicants and applicants' assignee, and that the Declaration and Power of Attorney to be filed hereafter will confirm the undersigned agent's authorization and appointment. Applicants are considered a small entity and assignee Deutsches Krebsforschungszentrum is also considered a small entity within the meaning of 37 CFR § 1.9.

				CALCULATIONS	PTO USE ONLY
				<b>09/914549</b>	
<p>17. <input checked="" type="checkbox"/> The following fees are submitted:</p> <p><b>Basic National Fee</b> (37 CFR 1.492(a)(1)-(5)): Search Report has been prepared by the EPO or JPO ..... \$860.00</p> <p>International preliminary examination fee paid to USPTO (37 CFR 1.482) ..... \$0.00</p> <p>No International preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2)) ..... \$0.00</p> <p>Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO ..... \$1000.00</p> <p>International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4) ..... \$0.00</p>					
<b>ENTER APPROPRIATE BASIC FEE AMOUNT =</b>				\$ 860.00	
<p>Surcharge of <b>\$130.00</b> for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(e)).</p>				\$	
Claims	Number Filed	Number Extra	Rate		
Total Claims	38-20 =	18	X \$18.00	\$ 324.00	
Independent Claims	12- 3 =	9	X \$80.00	\$ 720.00	
Multiple dependent claim(s) (if applicable)				+ \$000.00	\$
<b>TOTAL OF ABOVE CALCULATIONS =</b>				1904.00	
<p>Reduction by 1/2 for filing by small entity, if applicable. Verified Small Entity statement must also be filed. (Note 37 CFR 1.9, 1.27, 1.28).</p>				\$ 952.00	
<b>SUBTOTAL =</b>				\$ 952.00	
<p>Processing fee of <b>\$130.00</b> for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 Months from the earliest claimed priority date (37 CFR 1.492(f)).</p>				+	\$
<b>TOTAL NATIONAL FEE =</b>				\$ 952.00	
<p>Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). <b>\$40.00</b> per property</p>				+	\$
<b>TOTAL FEE ENCLOSED =</b>				\$ 430.00	
				Amount to be:	\$
				refunded	
				Charged	\$522.00
<p>a. <input checked="" type="checkbox"/> A check in the amount of <u>\$430.00</u> to cover part of the above fees is enclosed.</p> <p>b. <input checked="" type="checkbox"/> Please charge my Deposit Account No. <u>08-3284</u> in the amount of <u>\$522.00</u> to cover the above fees. A duplicate copy of this sheet is enclosed.</p> <p>c. <input checked="" type="checkbox"/> The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. <u>08-3284</u>. A duplicate copy of this sheet is enclosed.</p>					
<p><b>NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not yet been met, a petition to revive (37 CFR 1.127(a) or (b)) must be filed and granted to restore the application to pending status.</b></p>					
<p>SEND ALL CORRESPONDENCE TO:</p> <p><b>Steven J. Hultquist</b> <b>Intellectual Property/Technology Law</b> <b>P. O. Box 14329</b> <b>Research Triangle Park, NC 27709</b></p>					
 <b>MARIANNE FUERER</b> <b>Registration No. 39,983</b>					

Rec'd PCT/PTO 04 JAN 2002

4121-129  
PATENT APPLICATION

---

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

---

**In re Application of:** Poustka, et al.

**Application No.:** New U.S. National Stage Application of  
PCT International Application No.  
PCT/DE00/00583

**International Filing Date:** 28 February 2000

**Priority Date Claimed:** 26 February 1999 (German Appl. No. 199 048  
423.8)

**U.S. National Phase Filing Date:** Date of mailing identified below

**Title:** **PROTEIN (TP) THAT IS INVOLVED IN  
THE DEVELOPMENT OF THE NERVOUS  
SYSTEM**

---

**EXPRESS MAIL CERTIFICATE**

I hereby certify that I am mailing the attached documents to the  
Commissioner for Patents on the date specified, in an envelope  
addressed to the Commissioner for Patents, Box Patent Application,  
Washington, DC 20231, and Express Mailed under the provisions of  
37 CFR 1.10.

Blake Crouch

Name of Person Mailing This Document



Signature

August 24, 2001

Date

EL831358276US

Express Mail Label Number

---

**PRELIMINARY AMENDMENT**

---

Commissioner for Patents  
BOX PATENT APPLICATION  
Washington, D.C. 20231

Sir:

Prior to examination of the above-identified new national phase patent application, please amend the application, as follows:

**In the Claims**

Please amend claims 1-38 to read as follows:

1. A DNA sequence coding for a protein which is involved in the development of the nervous system, in particular the CNS, and is expressed in a tissue-specific and development-specific manner, wherein the DNA sequence comprises the following DNA sequences:
  - (a) the DNA sequence of figure 1, figure 2, figure 3, figure 4, figure 5, figure 6, figure 7 or figure 8;
  - (b) the DNA sequence of figure 9 or figure 10;
  - (c) the DNA sequence of figure 11;
  - (d) the DNA sequence of figure 12 or figure 13;
  - (e) the DNA sequence of figure 14 or figure 15;
  - (f) the DNA sequence of figure 16;
  - (g) the DNA sequence of figure 17 or 18;
  - (h) the DNA sequence of figure 19;
  - (i) a DNA sequence hybridizing with (a), (b), (c), (d), (e), (f), (g) or (h)
  - (j) fragments, variants, functional equivalents, derivatives or precursors of the

DNA sequence of (a), (b), (c), (d), (e), (f), (g), (h) or (i); or

(k) a DNA sequence which differs from the DNA sequence of (a), (b), (c), (d), (e), (f), (g), (h), (i) or (j) due to the degeneration of the genetic code.

2. The DNA sequence according to claim 1, which codes for a protein or peptide comprising the amino acid sequence of figure 1, figure 9, figure 11, figure 12, figure 13, figure 14, figure 15, figure 16, figure 17, figure 18 or figure 19, wherein the protein or peptide has the biological activity defined in claim 1.
3. An antisense RNA, characterized in that it is complementary to the DNA sequence of claim 1 and can reduce or inhibit the synthesis of the protein encoded by this DNA sequence.
4. A ribozyme, characterized in that it is complementary to the DNA sequence of claim 1 and can bind specifically to the RNA transcribed by this DNA sequence and can cleave it so as to reduce or inhibit the synthesis of the protein encoded by this DNA sequence.
5. An expression vector, containing the DNA sequence selected from the group consisting of the protein according to claim 1 the antisense RNA according to claim 3 or the ribozyme according to claim 4.
6. The expression vector according to claim 5, which comprises additionally the promoter of the human T gene or an ortholog of the T gene.
7. An expression vector according to claim 6, which codes for a protein selected from the group consisting of T, T2, T3 proteins or for fragments thereof in the form of a reporter fusion protein.
8. A host cell which is transformed with an expression vector selected from the

group consisting of the expression vector of claim 5, claim 6 and claim 7.

9. A protein which is encoded by the DNA sequence according to claim 1 and which is involved in the development of the nervous system and is expressed in tissue-specific and development-specific manner, or fusion proteins, fragments, variants, derivatives or precursors of the protein
10. Protein according to claim 9, which has one of the following motives:  
Motive 1:  
(A,T) (I,P,V) (L,T) (G,A,Q) (L,V)XXX(L,V)  
Motive 2:  
IYTDQWAN  
Motive 3:  
AXXXXXXXXXXGXXXXXXAXXXXXXXXXXXXXXXXXXXXXQ  
Motive 4:  
SXXXXDX (12,20) KX (17, 22)AXXXXXXXXXL  
Motive 5:  
IYTDWANXXLX (K, R)  
Motive 6:  
KX(18,21)AXXXXXXXXXLX(15,24) S  
Motive 7:  
NX (3,11) SXXXAXXXXXXXL  
wherein        X     every amino acid  
                    (A,T) = amino acid A or T at this site  
                    X(number 1, number 2) = number 1 to number 2  
                    Xs at this site.
11. A method of producing the protein according to claim 9, which comprises culturing the host cell according to claim 8 under suitable conditions and obtaining the protein from the cell or the culture medium.
12. Antibody which is directed against the protein according to claim 9 or fragment

thereof.

13. Antibody according to claim 12, which is obtained by immunizing animals with a peptide having the sequence "EKGEDPETRRMRTVKNIADI".
14. A method for preventing or treating diseases of the nervous system by using a member selected from the group consisting of the DNA sequence according to claim 1, the antisense RNA according to claim 3, the ribozyme according to claim 4, the expression vector according to any of claims 5 to 7, the protein according to claim 9 and the antibody or the fragment thereof according to claim 12 or 13 for preventing or treating diseases of the nervous system, in particular of the CNS.
15. The method according to claim 14, wherein the disease of the nervous system is a tumoral disease of the CNS.
16. The method according to claim 14, wherein the treatment of diseases of the nervous system are the promotion of the neuronal regeneration in the case of injuries of the nervous system and degenerative diseases of the nervous system.
17. The method according to claim 14, wherein the treatment of diseases of the nervous system are the regeneration of the neuronal linkages and the regeneration of the innate and acquired malfunctions of the nervous system.
18. The method according to claim 15 for inhibiting the growth and spreading of tumor cells.
19. Diagnostic method for detecting a disturbed expression of the protein according to claim 9 or for detecting a changed form of this protein, in which a sample is contacted with a member selected from the group consisting of the DNA sequence according to claim 1, the DNA sequence according to claim 2, the

antibody or the fragment thereof according to claim 12, and the antibody or the fragment thereof of claim 13 and then it is determined directly or indirectly whether the concentration of the protein and/or its amino acid sequence differs from a protein obtained from a healthy patient.

20. Diagnostic kit for carrying out the method according to claim 19, which contains at least one member selected from the group consisting of the DNA sequence according to claim 1, the DNA sequence according to claim 2, the antibody or the fragment thereof according to claim 12, and the antibody or the fragment thereof according to claim 13.
21. Non-human mammal whose naturally occurring T, T2 or T3 gene comprises a change in the gene structure or the gene sequence.
22. Non-human mammal, wherein a change of the gene structure of the T, T2 or G3 gene is achieved in the mammal by introducing a deletion in place of which a homologous or heterologous sequence is introduced.
23. Non-human mammal, wherein a change of the gene structure of the T, T2 or C3 gene is achieved by inserting a homologous or heterologous sequence in the corresponding gene naturally occurring in the mammal.
24. Non-human mammal according to claim 22, wherein the heterologous sequence is the selection marker sequence.
25. Non-human mammal according to claim 24, wherein the selection marker sequence conveys resistance to neomycin.
26. A method of producing a non-human mammal selected from the group consisting of the non-human mammal according to claim 21, claim 22, claim 23, claim 24 and claim 25, characterized by the steps of:
  - (a) producing a DNA fragment, in particular a vector, containing a changed

T, T2 or G3 gene, the T, T2 or T3 gene having been modified by inserting a heterologous sequence, in particular a selectable marker;

- (b) preparing embryonal stem cells from a non-human mammal (preferably a mouse);
- (c) transforming the embryonal stem cells from step (b) with the DNA fragment from step (a), the T gene in the embryonal stem cells being changed by homologous recombination with the DNA fragment from (a)
- (d) culturing the cells from step (c),
- (e) selecting the cultured cells from step (d) for the presence of the heterologous sequence, in particular the selectable marker,
- (f) producing chimeric non-human mammals from the cells of step (e) by injecting these cells into mammalian blastocysts (preferably mouse blastocysts), transferring the blastocysts to pseudo-pregnant female mammals (preferably mouse) and analyzing the resulting offspring for a change of the T, T2 or T3 gene.

27. Transgenic cell or tissue which is capable of expressing a T protein or part of the T protein or an ortholog thereof.

28. A method for the analysis of the function of the T gene family by using a member selected from the group consisting of the non-human mammal according to claim 21, claim 22, claim 23, claim 24, claim 25 the transgenic cell of claim 27 or the transgenic tissue according to claim 27.

29. A method for identifying inhibitors and enhancers of the T gene family by using the non-human mammal according to claim 21, claim 22, claim 23, claim 24, claim 25, the transgenic cell according to claim 27 or the transgenic tissue according to claim 27.

30. Vertebrate gene and functional equivalent, derivative or a bioprecursor thereof, which code for a protein having a statistically significant amino acid sequence homology to the T gene, T2 gene or T3 gene according to any of the following

figures selected from the group consisting of: figure 1, figure 9, figure 11, figure 12, figure 13, figure 14, figure 15, figure 16, figure 17, figure 18 or figure 19.

31. T gene and its vertebrate orthologs and vertebrate paralogs which code for a nuclear pore protein.
32. Vertebrate protein which has an amino acid sequence according to figure 1 or an amino acid sequence which differs from the amino acid sequence in figure 1 by one or more amino acids.
33. Vertebrate T, T2 or T3 gene and the protein encoded therein in all of its naturally occurring allelic and mutated forms.
34. Medicament containing a protein according to claim 9 or a functional equivalent, a fragment or a bioprecursor thereof in combination with a pharmaceutically acceptable carrier.
35. The method of identifying substances which has an enhancing or inhibiting influence on the effect of T protein, T2 protein or T3 protein, by means of determining the bi-directional transport through the nuclear pores,  
determining the binding to filaments of the cell (e.g. actin filaments and microtubuli) or  
determining the increased or reduced transcription of cellular or reporter genes.
36. Method of identifying substances which have an enhancing or inhibiting influence on the effect of proteins which are functionally linked to the T protein in direct or indirect way, or represent parallel signal or functional pathways, by means of  
- determining the bi-directional transport through the nuclear pores,

- determining the phosphorylation and the dephosphorylation of proteins,
- determining the binding of the T protein to filaments of the cell (e.g. actin filaments and microtubuli), or
- determining the increased or reduced transcription of cellular or reporter genes.

37. The method according to claim 35, wherein the modified transcription with reporter molecules, preferably the occurrence of certain mRNAs or the EGEP protein, is detected.

38. The method of identifying further proteins which play a role in the development and function of the nervous system and/or are a nuclear pore protein, wherein the method comprises the steps of:

- (a) producing an antibody against a protein according to claim 9,
- (b) contacting a cell extract with the antibody and identifying the antibody/protein complex,
- (c) analyzing the complex to identify a protein which has bound to the protein of the complex and is no antibody, and
- (d) optionally repeating steps (a) to (c) to identify further proteins of this function.

**REMARKS**

A marked-up version of amended paragraph in the specification and amended claims 1-38 are included herewith in Appendix A.

It is requested that the examination and prosecution of this application proceed on the basis of the English translation of the PCT International application included herewith and these amended claims 1-38.

Respectfully submitted,



Marianne Fuierer  
Registration No. 39,983  
Attorney for Applicants

INTELLECTUAL PROPERTY/  
TECHNOLOGY LAW  
P. O. Box 14329  
Research Triangle Park, NC 27709  
Phone: (919) 419-9350  
Fax: (919) 419-9354  
Attorney File: 4121-129

**APPENDIX A**

1. A DNA sequence coding for a protein which is involved in the development of the nervous system, in particular the CNS, and is expressed in a tissue-specific and development-specific manner, wherein the DNA sequence comprises the following DNA sequences:

- (a) the DNA sequence of figure 1, figure 2, figure 3, figure 4, figure 5, figure 6, figure 7 or figure 8;
- (b) the DNA sequence of figure 9 or figure 10;
- (c) the DNA sequence of figure 11;
- (d) the DNA sequence of figure 12 or figure 13;
- (e) the DNA sequence of figure 14 or figure 15;
- (f) the DNA sequence of figure 16;
- (g) the DNA sequence of figure 17 or 18;
- (h) the DNA sequence of figure 19;
- (i) a DNA sequence hybridizing with (a), (b), (c), (d), (e), (f), (g) or (h)
- (j) fragments, variants, functional equivalents, derivatives or precursors of the DNA sequence of (a), (b), (c), (d), (e), (f), (g), (h) or (i); or
- (k) a DNA sequence which differs from the DNA sequence of (a), (b), (c), (d), (e), (f), (g), (h), (i) or (j) due to the degeneration of the genetic code.

3. An antisense RNA, characterized in that it is complementary to the DNA sequence of claim 1 [or 2] and can reduce or inhibit the synthesis of the protein encoded by this DNA sequence.
4. A ribozyme [Ribozyme], characterized in that it is complementary to the DNA sequence of claim 1 [or 2] and can bind specifically to the RNA transcribed by this DNA sequence and can cleave it so as to reduce or inhibit the synthesis of the protein encoded by this DNA sequence.
5. An expression [Expression] vector, containing the DNA sequence selected from the group consisting of the protein according to claim 1 [or 2 or coding for] the antisense RNA according to claim 3 or the ribozyme according to claim 4.
7. An expression [Expression] vector according to claim [5 or] 6, which codes for a protein selected from the group consisting of [for the] T, T2, [or] T3 proteins or for fragments thereof in the form of a reporter fusion protein.
8. A host [Host] cell which is transformed with [the] an expression vector selected from the group consisting of the expression vector of claim 5, claim 6 and claim 7.[according to any of claims 5 to 7.]
9. A protein [Protein] which is encoded by the DNA sequence according to claim 1 [or 2] and which is involved in the development of the nervous system and is expressed in tissue-specific and development-specific manner, or fusion proteins, fragments, variants, derivatives or precursors of the protein
11. A method [Method] of producing the protein according to claim 9, which comprises culturing the host cell according to claim 8 under suitable conditions and obtaining the protein from the cell or the culture medium.

14. A method for preventing or treating diseases of the nervous system by using a member selected from the group consisting of [Use of ]the DNA sequence according to claim 1 [or 2], the antisense RNA according to claim 3, the ribozyme according to claim 4, the expression vector according to any of claims 5 to 7, the protein according to claim 9 and [or] the antibody or the fragment thereof according to claim 12 or 13 for preventing or treating diseases of the nervous system, in particular of the CNS.
15. The method [Use] according to claim 14, wherein the disease of the nervous system is a tumoral disease of the CNS.
16. The method [Use] according to claim 14, wherein the treatment of diseases of the nervous system are the promotion of the neuronal regeneration in the case of injuries of the nervous system and degenerative diseases of the nervous system.
17. The method [Use] according to claim 14, wherein the treatment of diseases of the nervous system are the regeneration of the neuronal linkages and the regeneration of the innate and acquired malfunctions of the nervous system.
18. The method [Use] according to claim 15 for inhibiting the growth and spreading of tumor cells.
19. Diagnostic method for detecting a disturbed expression of the protein according to claim 9 or for detecting a changed form of this protein, in which a sample is contacted with a member selected from the group consisting of the DNA sequence according to claim 1. the DNA sequence according to claim 2, [or 2 or] the antibody or the fragment thereof according to claim 12, and the antibody or the fragment thereof of claim [or] 13 and then it is determined directly or indirectly whether the concentration of the protein and/or its amino acid sequence differs from a protein obtained from a healthy patient.

20. Diagnostic kit for carrying out the method according to claim 19, which contains at least one member selected from the group consisting of the DNA sequence according to claim 1, the DNA sequence according to claim [or] 2, [and/or] the antibody or the fragment thereof according to claim 12, and the antibody or the fragment thereof according to claim [or] 13.
24. Non-human mammal according to claim 22 [or 23], wherein the heterologous sequence is the selection marker sequence.
26. A method of producing a non-human mammal selected from the group consisting of the non-human mammal according to claim 21, claim 22, claim 23, claim 24 and claim 25,[ to any of claims 21 to 25], characterized by the steps of:
  - (a) producing a DNA fragment, in particular a vector, containing a changed T, T2 or G3 gene, the T, T2 or T3 gene having been modified by inserting a heterologous sequence, in particular a selectable marker;
  - (b) preparing embryonal stem cells from a non-human mammal (preferably a mouse);
  - (c) transforming the embryonal stem cells from step (b) with the DNA fragment from step (a), the T gene in the embryonal stem cells being changed by homologous recombination with the DNA fragment from (a)
  - (d) culturing the cells from step (c),
  - (e) selecting the cultured cells from step (d) for the presence of the heterologous sequence, in particular the selectable marker,
  - (f) producing chimeric non-human mammals from the cells of step (e) by injecting these cells into mammalian blastocysts (preferably mouse blastocysts), transferring the blastocysts to pseudo-pregnant female mammals (preferably mouse) and analyzing the resulting offspring for a change of the T, T2 or T3 gene.
28. A method for the analysis of the function of the T gene family by using a member selected from the group consisting of the [Use of the] non-human mammal

according to [any of claims] claim 21, claim 22, claim 23, claim 24, claim [to] 25[ or] the transgenic cell of claim 27 or the transgenic tissue according to claim 27. [for the analysis of the function of the T gene family.]

29. A method for identifying inhibitors and enhancers of the T gene family by using [Use of] the non-human mammal according to claim 21, claim 22, claim 23, claim 24, claim 25, [to any of claims 21 to 25 or] the transgenic cell according to claim 27 or the transgenic tissue according to claim 27. [for identifying inhibitors and enhancers of the T gene family.]
30. Vertebrate gene and functional equivalent, derivative or a bioprecursor thereof, which code for a protein having a statistically significant amino acid sequence homology to the T gene, T2 gene or T3 gene according to any of the following figures selected from the group consisting of: figure 1, figure 9, figure 11, figure 12, figure 13, figure 14, figure 15, figure 16, figure 17, figure 18 or figure 19.
37. The method according to claim 35 [or 36], wherein the modified transcription with reporter molecules, preferably the occurrence of certain mRNAs or the EGEP protein, is detected.



JC05 Rec'd PCT/PTO , 04 JAN 2002

4121-129

PATENT APPLICATION

*REC'D  
SJS  
HJS*

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Poustka, et al.



Application No.:

09/914,549

23448

PATENT TRADEMARK OFFICE

International Application No.:

PCT/DE00/00583

Priority Date Claimed:

28 February 2000 and 26 February 1999  
(German Appl. No. 199 048 423.8)

Title:

**PROTEIN (TP) THAT IS INVOLVED IN THE  
DEVELOPMENT OF THE NERVOUS  
SYSTEM**

FIRST CLASS MAIL CERTIFICATE

I hereby certify that I am mailing the attached documents to the Commissioner for Patents on the date specified, in an envelope addressed to the Commissioner for Patents, Washington, DC 20231, and First Class Mailed under the provisions of 37 CFR 1.8.

*Lee Ann Brown*

Lee Ann Brown

November 14, 2001

Date of Mailing

SECOND SUPPLEMENTAL PRELIMINARY AMENDMENT IN U.S. PATENT  
APPLICATION NO. 09/914,549

Commissioner for Patents  
Washington, D.C. 20231

Sir:

Prior to examination of the above-identified national phase patent application, please amend the application, as follows:

**In the Specification**

Please insert on page 1 between the title of the application and the first paragraph the following new paragraph:

**CROSS-REFERENCE TO RELATED APPLICATIONS**

This application is filed under the provisions of 35 U. S.C. §371 and claims the priority of International Patent Application No. PCT/DE00/00583 filed February 28, 2000, which in turn claims priority of German Patent Application No. 199 048 423.8 filed on February 26, 1999.

**REMARKS**

This claim to priority is being filed before the above-identified application meets all the requirements under 35 U.S.C. §371(b).

Respectfully submitted,



Marianne Fuerer  
Registration No. 39,983  
Attorney for Applicants

INTELLECTUAL PROPERTY/  
TECHNOLOGY LAW  
P. O. Box 14329  
Research Triangle Park, NC 27709  
Phone: (919) 419-9350  
Fax: (919) 419-9354  
Attorney File: 4121-129

09/914549

518 Rec'd PCT/PTO 24 AUG 2001

K 3008

124/pbk

Protein (TP) That is Involved in the Development of the Nervous System

The present invention relates to a protein (T protein) and to proteins related thereto which are involved in the development of the nervous system and are expressed in a tissue-specific and development-specific manner, to the below described variants of these proteins and to DNA sequences coding for these proteins. The present invention further relates to antibodies directed against these proteins or to fragments thereof as well as to antisense RNAs or ribozymes directed against the expression of these proteins. Finally, the present invention concerns medicaments and diagnostic methods in which the above-mentioned compounds are used.

Mutations in genes playing a part in the development and maintenance of the nervous system are of utmost scientific and economic significance, since diseases of the nervous system, in particular CNS, occur frequently, are often characterized by a severe, partly fatal disease process and can be treated only to a limited extent thus far. The increase in the life expectancy is accompanied by a drastic increase in neurological and psychic diseases. The latter greatly limit the quality of life of the affected persons and cause considerable costs for both the affected person and the public.

Isolating and analyzing genes specific to the nervous system offer a good possibility of studying diseases, such as schizophrenia, Alzheimer's disease, autism, manic depression

and mental retardation, and eventually of also being able to treat them.

The present invention is thus based on the technical problem of providing products by means of which disturbances in the development and function of the nervous system can be diagnosed and optionally be treated.

The solution to this technical problem is achieved by providing the embodiments characterized in the claims.

The subject matter of the present invention is thus a DNA sequence coding for a protein which is involved in the development and function of the nervous system, in particular the CNS, and is expressed in tissue-specific and development-specific manner, the DNA sequence comprising the following DNA sequences:

- (a) the DNA sequence of figure 1, figure 2, figure 3, figure 4, figure 5, figure 6, figure 7 or figure 8;
- (b) the DNA sequence of figure 9 or figure 10;
- (c) the DNA sequence of figure 11;
- (d) the DNA sequence of figure 12 or figure 13;
- (e) the DNA sequence of figure 14 or figure 15;
- (f) the DNA sequence of figure 16;
- (g) the DNA sequence of figure 17 or 18;
- (h) the DNA sequence of figure 19;
- (i) a DNA sequence hybridizing with (a), (b), (c), (d), (e), (f), (g) or (h);
- (j) variants, derivatives, precursors or fragments of the DNA sequence of (a), (b), (c), (d), (e), (f), (g), (h) or (i); or

(k) a DNA sequence differing from the DNA sequence of  
(a), (b), (c), (d), (e), (f), (g), (h), (i) or (j)  
due to the degeneration of the genetic code.

The present invention is based on the isolation of a human DNA sequence (referred to as gene "T" or T gene; see figures 1 to 8, which codes for protein TP), it turning out that the protein encoded by this DNA sequence is required in the nervous system. In this connection, the expression of the gene encoding this protein is increased in the nervous system. A sequence analysis showed that it is a new gene. Moreover, further genes could be isolated which have homologies to this gene (murine gene "T", figures 9 and 10; human gene "T2", figure 16; human gene "T3", figures 17 and 18; murine gene T2, figures 12 and 13; murine gene T3, figure 19). The T gene, T2 gene and T3 gene are members of the T (gene) family, as shown below, and originate preferably from vertebrates, such as man, mouse or rat. Defects in these genes limit the functions of the nervous system, in particular the CNS. These genes also perform an important function in the control of cell growth, and changes in these genes or their expression result in defects regarding the control of cell growth, e.g. also in tumor formation, in particular of the neuroblastoma. Small children up to the age of 8 are affected almost exclusively by this cancerous disease. The first symptoms already occur within the first 12 months of life in 25 to 30 percent of the cases. In the case of the neuroblastoma very young cells of the autonomous nervous system degenerate. Since these nerves extend along the rear side of the abdominal region and the chest, neuroblastomas usually occur in the regions of the stomach, pelvis, chest and neck. More than half the diseases start from the suprarenal marrow which is also formed by nerve cells. Symptoms which may refer in small

children to a neuroblastoma are nodes, swellings, bone pain, limping, tiredness, fever, paleness, sweating, obstinate or persistent cough, hematomas around the eye. A neuroblastoma can be diagnosed by a physician by means of blood tests, urine analyses and ultrasonic examinations and by the removal of biopsies from the tumor and an examination of bone marrow. As soon as the accurate location of the tumor is diagnosed, it is removed by means of an operation. However, the early formation of metastases creates a problem. By isolating and analyzing the T gene it is now possible to develop novel measures of diagnosing and treating the neuroblastoma. Due to this, it is possible to diagnose the cancerous disease early and establish forms of therapy promising better chances of recovery.

Mutations in genes of the T gene family also lead to a disturbed development and differentiation of the nervous system, in particular the brain. In many cases, this results in mental diseases, e.g. mental retardations or Alzheimer's disease. The T gene also plays an important role in the interconnection of individual regions of the brain, e.g. forebrain and midbrain. Mutations in this gene lead in some cases to schizophrenic diseases and syndromes of autism. By means of the human and murine genes it is possible to draw important fundamental conclusions as to the development of the nervous system and in particular the brain. Good approaches offer themselves as regards the research of pathologic changes of the nervous system and in particular the brain.

Patients can be examined more simply for possible mutations by means of the genomic sequences. The genomic sequences of the T gene are of advantage in particular when little (tumor) material is available for the analysis. By this it

is possible, for example, to examine even minute tumors for mutations in this gene. This also provides the possibility of checking a therapy (in particular radiation therapy and/or chemotherapy) for its being successful, since it is possible to detect tumor cells circulating in the blood by genomic primers which are specific to the genomic DNA using a PCR reaction.

The term "hybridizing" used in the present invention relates to conventional hybridization conditions, preferably to hybridization conditions which use 5xSSPE, 1 % SDS, 1xDenhardt's solution as the solution and where hybridization temperatures are between 35°C and 70°C, preferably 65°C. Following hybridization, washing is preferably carried out using first 2xSSC, 1 % SDS and then 0.2xSSC at temperatures between 35°C and 70°C, preferably of 65°C (regarding a definition for SSPE, SSC and Denhardt's solution see Sambrook et al., Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)). Stringent hybridization conditions are particularly preferred, as described in Sambrook et al., *supra*, for example.

The terms "variants" or "fragment" used in the present invention comprise DNA sequences which differ from the sequences indicated in the figures by deletion(s), insertion(s), substitution(s) and/or other modifications known in the art or comprise a fragment of the original nucleic acid molecule, the protein or peptide encoded by these DNA sequences still having the above-mentioned properties. Therefore, functional equivalents, derivatives, precursors (bioprecursors) are counted among them. Derivatives are understood to mean e.g. mutation derivatives

(produced by deletions or insertions, for example), fusions, allele variants, muteins and splicing variants. Two select examples of such splicing variants are shown in figures 14 and 15. Methods of producing the above changes in the nucleic acid sequence are known to a person skilled in the art and are described in standard works of molecular biology, e.g. in Sambrook *et al.*, *supra*. The person skilled in the art is also capable of determining whether a protein encoded by a nucleic acid sequence modified in such a way still has the above-mentioned properties.

In a preferred embodiment, the present invention relates to a DNA sequence which encodes a protein comprising the amino acid sequence of figure 1, figure 9, figure 11, figure 12, figure 13, figure 14, figure 15, figure 16, figure 17, figure 18 or figure 19, the protein having the above-defined biological activity.

By lowering or inhibiting the expression of the above described DNA sequences it is possible to reduce or eliminate the synthesis of the proteins encoded by them, e.g. the T protein, which is desirable for certain states of a disease, for example. Therefore, another preferred embodiment of the present invention relates to antisense RNA, which is characterized in that it is complementary to the above DNA sequences and can reduce or inhibit the synthesis of the protein encoded by these DNA sequences and to a ribozyme, which is characterized in that it can bind specifically to part of the above DNA sequences and to the RNA transcribed by these DNA sequences and can cleave them so as to reduce or inhibit the synthesis of the protein encoded by these DNA sequences. These antisense RNAs and ribozymes are preferably complementary to a coding region of the mRNA. Based on the disclosed DNA sequences, the person

skilled in the art can produce and use suitable antisense RNAs. Suitable methods are described in EP-B1 0 223 399 or EP-A1 0 458, for example. Ribozymes are RNA enzymes and consist of a single RNA strand. They can cleave intermolecularly other RNAs, e.g. the mRNAs transcribed by the DNA sequences according to the invention. These ribozymes must, in principle, have two domains: (1) a catalytic domain and (2) a domain which is complementary to the target RNA and can bind thereto, which is a precondition for a cleavage of the target RNA. Based on the methods described in the literature, it is meanwhile possible to construct specific ribozymes which excise a desired RNA at a certain pre-select site (see e.g. Tanner et al., in: *Antisense Research and Applications*, CRC Press, Inc. (1993), 415-426).

The DNA sequences according to the invention or the DNAs encoding the above described antisense RNAs or ribozymes may also be inserted in a vector or expression vector. Thus, the present invention also comprises vectors or expression vectors containing these DNA sequences. The term "vector" relates to a plasmid (e.g. pUC18, pBR322, pBlueScript), to a virus or another suitable vehicle. In a preferred embodiment, the DNA molecule according to the invention is functionally linked in the vector to regulatory elements allowing the expression thereof in prokaryotic or eukaryotic host cells. Along with the regulatory elements, e.g. a promoter, such vectors contain typically a replication origin and specific genes which allow the phenotypic selection of a transformed host cell. The lac, trp promoter or the T7 promoter are counted among the regulatory elements for the expression in prokaryotes, e.g. *E. coli*, those for the expression in eukaryotes comprise the AOX1 or GAL1 promoter in yeast, and those for the expression in animal

cells include the CMV, SV40, RVS40 promoter, CMV or SV40 enhancer. Further examples of suitable promoters are the metallothionein I promoter and the polyhedrin promoter. In a preferred embodiment the vector contains the promoter of the human T gene or an ortholog of the T gene. Suitable expression vectors for *E. coli* are e.g. pGEMEX, pUC derivatives, pGEX-2T, pET3b and pQE-8, the latter being preferred. Suitable vectors for the expression in yeast comprise pY100 and Ycpad1, and suitable vectors for the expression in mammalian cells include pMSXND, pKCR, pEFBOS, cDM8 and pCEV4. Vectors derived from baculovirus for expression in insect cells, e.g. pAcSGHisNT-A, are also counted among the expression vectors according to the invention.

General methods known in the art can be used for constructing expression vectors which contain the DNA sequences according to the invention and suitable control sequences. These methods comprise e.g. *in vitro* recombination techniques, synthetic methods, and *in vivo* recombination techniques, as described in Sambrook *et al.*, *supra*, for example. The DNA sequences according to the invention can also be inserted in combination with a DNA coding for another protein or peptide, so that the DNA sequences according to the invention can be expressed in the form of a fusion protein, for example. These other DNAs are preferably reporter sequences which code for a reporter molecule comprising a detectable protein, e.g. a stain or coloring matter, an antibiotic resistance,  $\beta$ -galactosidase or a substances detectable by spectrophotometric, spectrofluorometric, luminescent or radioactive assays.

The present invention also relates to host cells containing the above described vectors. These host cells comprise

bacteria (e.g. the *E. coli* strains HB101, DH1, x1776, JM101, JM109, BL21 and SG13009), fungi, e.g. yeasts, preferably *S. cerevisiae*, plant cells, insect cells, preferably sf9 cells, and animal cells, preferably cells from vertebrates or mammals. Preferred mammalian cells are CHO, VERO, BHK, HeLa, COS, MDCK, 293 or WI38 cells. Methods of transforming these host cells for the phenotypic selection of transformants and for the expression of the DNA molecules according to the invention using the above-described vectors are known in the art.

The genes belonging to the sequences according to the invention can be amplified by suitable primer sequences. The primer sequences indicated in figure 20 are particularly suited for amplification of genes T2 and T3.

The present invention also relates to the proteins encoded by the DNA sequences according to the invention and to methods of producing the protein encoded by the DNA sequences according to the invention. The person skilled in the art is familiar with conditions of culturing transformed or transfected host cells. The method according to the invention comprises the culturing of the above described host cells under conditions which allow the expression of the protein (or fusion protein) (preferably stable expression) and the collection of the protein from the culture or from the host cells. Suitable purification methods (e.g. preparative chromatography, affinity chromatography, e.g. immunoaffinitychromatography, HPLC, etc.) are generally known.

The proteins according to the invention preferably comprise the amino acid sequences shown in figure 1, figure 9, figure 11, figure 12, figure 13, figure 14, figure 15, figure 16,

figure 17, figure 18 or figure 19 or represent fusions, fragments, derivatives or precursors (bioprecursors) thereof, the above mentioned properties being maintained within the meaning of functional equivalents. As to the definitions of these terms, reference is made to the respective explanations above. Derivatives are understood to mean in particular the changed proteins or peptides which differ from the sequences shown in the figures by conservative amino acid substitutions or contain non-conserved amino acid substitutions that do not change the function of the T proteins to a substantial degree.

The following amino acid motives have been identified by Inventors. They are suited to identify formerly unknown proteins which belong to the T/T2/T3 family according to the invention and a protein superfamily from pore membrane proteins and filament-binding proteins.

Motive 1:

(A, T) (I, P, V) (L, T) (G, A, Q) (L, V) XXX (L, V)

Motive 2:

IYTDWAN

Motive 3:

AXXXXXXXGXXXXXAXXXXXXXXXXXXXXXXXXXXXQ

Motive 4:

SXXXXDX(12,20)KX(17,22)AXXXXXXXL

Motive 5:

IYTDWANXXLX(K, R)

Motive 6:

KX(18,21)AXXXXXXXLX(15,24)S

Motive 7:

NX(3,11)SXXXAXXXXXXXL

Explanation: X stands for every amino acid

(A, T) means amino acid A or T at this site

X(2,4) denotes two to four Xs at this site

Another preferred embodiment of the present invention relates to antibodies against the above described proteins according to the invention or to a fragment thereof. These antibodies may be monoclonal, polyclonal or synthetic antibodies or fragments thereof. In this connection, the term "fragment" means all parts of the monoclonal antibody (e.g. Fab, Fv or "single chain Fv" fragments) which have an epitope specificity the same as that of the complete antibody. The person skilled in the art is familiar with the production of such fragments.

The antibodies according to the invention are preferably monoclonal antibodies. The antibodies according to the invention can be produced according to standard methods, the protein encoded by the DNA sequences according to the invention or a synthetic fragment thereof serving as an immunogen. Methods of obtaining monoclonal antibodies are known to the person skilled in the art and comprise e.g. as a first step the production of polyclonal antibodies using the proteins according to the invention or fragments thereof (synthetic peptides, for example) as an immunogen for immunizing suitable animals, e.g. rabbits or chickens, and the collection of the polyclonal antibodies from the serum or egg yolk.

For example, cell hybrids from cells producing antibodies and tumor cells from bone marrow are then produced and cloned. Thereafter, a clone is selected which produces an antibody specific to the antigen used. This antibody is then produced. Examples of cells producing antibodies are spleen cells, lymph node cells, B lymphocytes, etc. Examples of animals which can be immunized for this purpose are mice,

rats, horses, goats and rabbits. The myeloma cells can be obtained from mice, rats, humans or other sources. The cell fusion can be carried out by the generally known method developed by Köhler and Milstein, for example. The hybridomas obtained by cell fusion are screened using the antigen according to the enzyme-antibody method or according to a similar method. Clones are obtained with the boundary dilution method, for example. The resulting clones are implanted intraperitoneally into BALB/c mice, for example, the mouse ascites is removed after 10 to 14 days, and the monoclonal antibody is purified by known methods (e.g. ammonium sulfate fractionation, PEG fractionation, ion exchange chromatography, gel chromatography or affinity chromatography).

In a particularly preferred embodiment, said monoclonal antibody is an antibody originating from an animal (e.g. mouse), a humanized antibody or a chimeric antibody or a fragment thereof. Chimeric antibodies similar to human antibodies or humanized antibodies have a reduced potential antigenicity, however, their affinity is not lowered over the target. The production of chimeric and humanized antibodies or of antibodies similar to human antibodies has been described in detail (see e.g. Queen et al., Proc. Natl. Acad. Sci., U.S.A. 86 (1989), 10029, and Verhoeyan et al., Science, 239 (1988), 1534). Humanized immunoglobulins have variable framework regions which originate substantially from a human immunoglobulin (designated acceptor immunoglobulin) and the complementarity of the determining regions which originate substantially from a non-human immunoglobulin (e.g. from a mouse) (designated donor immunoglobulin). The constant region(s) originate(s), if available, also substantially from a human immunoglobulin. When administered to human patients, humanized (and the

human) antibodies have a number of advantages over antibodies from mice or other species: (a) the human immune system should not regard the framework or the constant region of the humanized antibody as foreign and therefore the antibody response to such an injected antibody should be less than to that to a completely foreign mouse antibody or a partially foreign chimeric antibody; (b) since the effector region of the humanized antibody is human, it might interact better with other parts of the human immune system, and (c) injected humanized antibodies have a half life which is substantially equivalent to that of human antibodies occurring in nature, which permits the administration of doses smaller and less frequent as compared to antibodies of other species.

The antibodies according to the invention can be used for the immunoprecipitation of the above discussed proteins, for the isolation of related proteins from cDNA expression libraries or for the below indicated purposes (diagnosis/therapy), for example.

The present invention also relates to a hybridoma which produces the above described monoclonal antibody.

In a preferred embodiment, the present invention relates to antibodies against the peptides of genes T2 and T3 listed separately (*cf.* figure 20).

It has been found that the below peptide can be used specifically for generating antibodies against the T protein. The amino acid sequence of the suitable peptide reads as follows:

EKGEDPETRRMRTVKNIAD

The present invention makes possible to study disturbances in the development and function of the nervous system on a genetic level. These disturbances comprise *inter alia* neurological and psychiatric diseases (*inter alia* Alzheimer's disease, Parkinson's disease, schizophrenia, manic-depressive diseases, autism, mental retardations), injuries of the nervous system, innate damage of the nervous system or degenerative diseases of the nervous system. The invention also enables the treatment of cancer, *inter alia* of tumors of the nervous system, such as neuroblastoma, astrocytoma, glioblastoma, medulloblastoma. This diagnosis cannot only be made postnatally but already prenatally. It can be detected by means of the DNA sequence according to the invention or probes or primers derived therefrom whether mammals, in particular humans, contain a gene which codes for and/or expresses the protein according to the invention or whether this gene results in a mutated form of the protein which is no longer biologically active. For this purpose, the person skilled in the art can carry out common methods, such as reverse transcription, PCR, LCR, hybridization and sequencing. The antibodies according to the invention are also suited e.g. for diagnosis, i.e. for detecting in a sample the presence and/or concentration of the protein according to the invention, a shortened or extended form of the protein, etc. The antibodies can be bound e.g. in immunoassays in liquid phase or to a solid carrier. In this case, the antibodies can be labeled in various ways. Suitable markers and labeling methods are known in the art. Examples of immunoassays are ELISA and RIA.

Thus, the present invention also relates to a diagnostic method for detecting a disturbed expression of the protein

according to the invention or for detecting a changed form of this protein, in which a sample is contacted with the DNA sequences according to the invention or the antibody according to the invention or the fragment thereof and then it is determined directly or indirectly whether the concentration of the protein and/or its amino acid sequence differs from a protein obtained from a healthy patient.

The present invention also allows to carry out therapeutic measures in connection with the above discussed disturbances, i.e. the above described inventive DNA sequences, antisense RNAs, ribozymes and antibodies can also be used for producing a medicament, e.g. for controlling the expression of the protein according to the invention, or for exchanging a mutated form of the gene by a functional form and thus also for the production of a medicament for preventing or treating diseases of the nervous system, in particular tumoral diseases of the CNS. For example, the protein according to the invention can be introduced into mammals, in particular humans, by common measures. For this purpose, it may be favorable to link the protein to a protein which is not considered foreign by the respective body, e.g. transferrin or bovine serum albumin (BSA). An inventive DNA sequence, antisense RNA or ribozyme can also be introduced into mammals, in particular humans, and expressed. By means of an antibody according to the invention it is possible to control and regulate the expression of the protein (TP) according to the invention or the related proteins.

Thus, the present invention also relates to a medicament which contains the above described DNA sequences, antisense RNA, the ribozyme, the expression vector, the protein according to the invention or the antibody or the fragment

thereof. This medicament contains, optionally in addition, a pharmaceutically compatible carrier. Suitable carriers and the formulation of such medicaments are known to the person skilled in the art. Suitable carriers are e.g. phosphate-buffered common salt solutions, water, emulsions, e.g. oil-in-water emulsions, wetting agents, sterile solutions, etc. The medicaments can be administered orally or parenterally. The topical, intra-arterial, intra-muscular, subcutaneous, intramedullary, intrathekal, intraventricular, intravenous, intraperitoneal or intranasal administration are counted among the methods for the parenteral administration. The suitable dose is determined by the attending physician and depends on various factors, e.g. on the age, sex and weight of the patient, the stage of the disease, the kind of administration, etc.

The above described nucleic acids are preferably inserted in a vector suitable for gene therapy and introduced into the cells under the control of a tissue-specific vector, for example. In a preferred embodiment, the vector containing the above described nucleic acids is a virus, e.g. an adenovirus, vaccinia virus or adenovirus. Retroviruses are particularly preferred. Examples of suitable retroviruses are MoMuLV, HaMuSV, MuMTV, RSV or GaLV. For the purposes of gene therapy, the nucleic acids according to the invention can also be transported to the target cells in the form of colloidal dispersions. They comprise liposomes or lipoplexes, for example (Mannino et al., Biotechniques 6 (1988), 682).

Finally the present invention relates to a diagnostic kit for carrying out the above described diagnostic method, which contains a DNA sequence according to the invention or the above described antibody according to the invention or a

fragment thereof. Depending on the kind of the kit, the DNA sequence or the antibody or the fragment thereof can be immobilized.

Sequences of the T genes can be applied to nylon membranes or glass carriers and hybridized with complex cDNA samples from tumors and pertinent normal tissues or diseased and pertinent healthy tissue. This enables the (fully automated) detection of the expression of these genes. The sequences used for this purpose can be e.g. the entire cDNA sequence or short sequence segments, e.g. 10-15 bp oligomers (see *inter alia* figure 20). Having determined the expression of the T genes, the therapy, *inter alia* the cancer therapy, can be selected deliberately according to the respective individual situation of the patient or can be adapted thereto. Genes whose changed expression influence already now the treatment of the patient are the N-myc gene in the case of neuroblastoma, for example. By detecting the expression of the T genes it is thus possible to adapt the treatment very quickly and efficiently to the respective requirements and in this way it contributes essentially to the improved therapy.

The isolation and characterization of the human gene according to the invention and in particular of the mouse homologues thereof also allow to establish an animal model, which is very valuable for the further study of diseases of the nervous system and of cancerous diseases on a molecular level. The subject matter of the present invention thus also relates to a non-human mammal whose T gene or T2 or T3 gene is changed, e.g. by inserting a heterologous sequence, in particular a selection marker sequence.

The expression "non-human mammal" comprises any mammal whose T gene or T2 or T3 gene can be changed. Examples of such mammals are mouse, rat, rabbit, horse, cattle, sheep, goat, monkey or ape, pig, dog and cat, with mouse being preferred.

The expression "T gene or T2 or T3 gene which is changed" signifies that a change of the gene structure or the gene sequence is carried out by standard methods in the corresponding gene occurring naturally in the non-human mammal. This can be achieved *inter alia* by introducing a deletion of about 1-2 kb, at the place of which a heterologous sequence, e.g. a construct for mediating antibiotic resistance (e.g. a "neo cassette") is introduced. Heterologous sequences allowing to carry out time-specific and tissue-specific deletions *in vivo* can also be inserted in the T gene. Furthermore, heterologous sequences allowing to track the expression of the T gene *in vivo* can be introduced into the T gene. This can be done *inter alia* by inserting a sequence coding for the GFP (green fluorescent protein) protein inside an exon or as an independent exon. These methods are generally described by Schwartzberg et al., Proc. Natl. Acad. Sci., U.S.A., Vol. 87, pages 3210-3214, 1990, to which reference is made herein.

In particular, the modification can be described and carried out as follows. Figure 9 represents part of the cDNA sequence of the T gene of a mouse. Illustration 10 shows an intron sequence of the T gene of a mouse, which is flanked by two exons. These murine sequences can then be used for the deliberate change of the murine T gene. For example, the splicing sequences of the intron can be deleted or changed such that the T gene is no longer spliced correctly. By incorporating a splicing acceptor sequence of another exon of the murine T gene into the intron sequence it is possible

to insert in this intron a sequence which is recognized as exon and is spliced to the T gene exon upstream thereof. This inserted sequence may be an exon, for example, which encodes the EGFP protein (EnhancedGreenFluorescentProtein). As a result, the original murine T gene becomes a fusion protein comprising the EGFP protein. Thus, a mouse can preferably be generated, which allows to track the expression of the T gene *in vivo*. The inserted sequence can be designed at its end (e.g. PolyA signal, splicing signals, etc.) such that no further exons of the T gene are spliced to the inserted exon or the spliced exon can no longer be translated. As a result, a deletion of the murine T protein forms on the C-terminal end or a premature discontinuance of the reading frame, and an (at least partial) inactivation of the protein function of the murine T gene can be achieved. It is also possible to insert, as new exon sequences, sequences which yield an mRNA sequence where this new mRNA sequence is localized at the 3' end. By suitable sequences it is then possible to achieve a change in the stability of the mRNA or a changed localization in the cell. The accompanying phenotypes of the thus modified mice can then result in important conclusions drawn on the function of the T gene. These mice can then also be used for detecting new active substances compensating the functional loss of the T gene.

In another preferred embodiment, the sequence of figure 13 is used for the production of a knock-out mouse. Figure 13 describes a murine sequence of gene T2. The elimination of the murine T2 genes can in this connection be achieved in different ways. For example, the splicing sequence (GT, underlined in figure 13) can be deleted or changed such that the T2 gene is no longer spliced correctly. In addition, by incorporating a splicing acceptor sequence of another exon

of the murine T2 gene into the following intron sequence it is possible to insert in this intron a sequence which is detected as exon and spliced to the T2 gene exons upstream thereof. This inserted exon may be e.g. an exon which codes for the EGFP protein. Due to this, the original murine T2 gene becomes a fusion protein which carries the EGFP protein at the C terminus. In this way, a mouse can be generated which allows to track the expression of the T2 gene *in vivo*. The inserted sequence can be designed at its end (e.g. PolyA signal, etc.) such that no further exons are spliced to the inserted exon by the T2 gene. A deletion of the murine T2 protein forms at the C-terminal end and an (at least partial) inactivation of the protein function of the murine T2 gene can be achieved. Such sequences can also be inserted as new exon sequences which yield an mRNA sequence in which at the 3' end this new mRNA sequence is localized. By means of suitable sequences it is then possible to achieve a change in the stability of the mRNA or a changed localization in the cell. The accompanying phenotypes of the thus changed mice can then lead to important conclusions as to the function of the T2 gene. These mice can also be used for detecting new active substances which compensate the functional loss of the T gene.

Furthermore, a mammal can be generated comprising a change in the T3 gene. The sequence in figure 19 represents part of the murine cDNA sequence of the T3 gene. Deliberate changes in the T3 gene of a mouse can be achieved by deletions or insertions. The inserted sequence can be an exon, for example, which codes for the EGFP protein. As a result, the original murine T3 gene becomes a fusion protein which carries the EGFP protein at the C terminus. Thus, a mouse can be generated which allows to track the expression of the T3 gene *in vivo*. The inserted sequence can be designed at

its end (e.g. PolyA signal, etc.) such that no further exons are spliced from the T3 gene to the inserted exon. A deletion of the murine T3 protein thus forms on the C-terminal end and an (at least partial) inactivation of the protein function of the murine T3 gene can be achieved. It is also possible to insert, as new exon sequences, sequences which yield an mRNA sequence where this new mRNA sequence is localized at the 3' end. By suitable sequences it is then possible to achieve a change in the stability of the mRNA or a changed localization in the cell. The accompanying phenotypes of the mice changed in this way can then lead to important conclusions as to the function of the T3 gene. These mice can then also be used for discovering new active substances which compensate the functional loss of the T3 gene.

Another subject matter of the present invention are cells which are obtained from the above non-human mammal. These cells can be present in any form, e.g. in a primary or long-term culture.

A non-human mammal according to the invention can be provided by common methods. A method is favorable which comprises the steps of:

- (a) producing a DNA fragment, in particular a vector, containing a changed T, T2 or G3 gene, the gene having been modified by inserting a heterologous sequence, in particular a selectable marker;
- (b) preparing embryonal stem cells from a non-human mammal (preferably a mouse);

- (c) transforming the embryonal stem cells from step (b) with the DNA fragment from step (a), the T gene in the embryonal stem cells being changed by homologous recombination with the DNA fragment from (a);
- (d) culturing the cells from step (c);
- (e) selecting the cultured cells from step (d) for the presence of the heterologous sequence, in particular the selectable marker;
- (f) producing chimeric non-human mammals from the cells from step (e) by injecting these cells into mammalian blastocysts (preferably mouse blastocysts), transferring the blastocysts to pseudo-pregnant female mammals (preferably mouse) and analyzing the resulting offspring for a change of the T gene.

In step (c), the mechanism of homologous recombination (*cf.* R.M. Torres, R. Kühn, *Laboratory Protocols for Conditional Gene Targeting*, Oxford University Press, 1997) is used to transfect embryonal stem cells. The homologous recombination between the DNA sequences present in a chromosome and new added cloned DNA sequences enable the insertion of a cloned gene in the genome of a living cell in place of the original gene. Using embryonal germ cells, animals which are homozygous for the desired gene or the desired gene portion or the desired mutation can be obtained via chimeras by this method.

The expression "embryonal stem cells" comprises any embryonal stem cells of a non-human mammal, suited to mutate the T gene. The embryonal stem cells originate preferably from a mouse, in particular the cells E14/1 or 129/SV.

The expression "vector" comprises any vector which by recombination with the DNA of embryonal stem cells enables a change of the T, T2 or G3 gene. The vector preferably has a marker by means of which selection can be made for existing stem cells in which the desired recombination has been made. Such a marker is the loxP/tk neo cassette, for example, which can be removed by means of the Cre/loxP system from the genome again.

The person skilled in the art also knows conditions and materials serving for carrying out steps (a) - (f).

By means of the present invention a non-human mammal is provided whose T, T2 or T3 gene is changed. This change may be an elimination of the gene expression-regulating function. Using such a mammal or cells therefrom it is possible to study selectively the gene expression-controlling function of the TP protein. It is also possible by this to find substances, medicaments and therapy approaches by which selective influence on the controlled function is possible. The present invention therefore provides a basis for influencing the most differing diseases. Such diseases are e.g. limitations of the CNS functions which cover mental retardations or the induction of cancer resulting from deficiencies in the control of cell proliferation.

Inventors found out in the sequence analysis that the T2 gene in the coding region of the cDNA sequence contains CGG trinucleotides which are known to be sensitive to methylation. The T2 gene thus has in the coding region (N-terminal region of the protein which has no homology to the T protein or T3 protein) a methylation-sensitive and

unstable sequence which results in the failure of the gene accompanied by a mental retardation and uncontrolled cell growth, such as cancer.

All the three genes of the T family play a major role in the case of tumors. The T gene is affected in many tumors by genomic rearrangements. For example, in neuroblastomas genomic changes in the DNA of tumors can be found as compared to the DNA of the accompanying healthy tissue. The expression of the T gene, e.g. in tumors of the brain, is also changed. A strongly changed expression can be found *inter alia* in the advanced stages of glioblastomas. Tumor-specific changes of the expression of the T gene and the occurrence of the T protein can also be detected in meningiomas.

In many tumoral diseases, the T2 gene also undergoes genomic rearrangements, and a changed expression can be detected in tumors. For example, in melanomas and lung tumors genomic rearrangements of the T2 gene can be detected. Expression differences are also detectable in gliomas, glioblastomas, astrocytomas and PNETs (Primitive Neuro-Ectodermal Tumors), for example.

In many tumors, the T3 gene also undergoes genomic rearrangements and expression changes. Rearrangements can be detected in colon carcinomas, for example. Expression differences are detectable *inter alia* in gliomas, glioblastomas, astrocytomas and PNETs (Primitive Neuro-Ectodermal Tumors).

By isolating and accurately analyzing the T gene, Inventors now have found that the T protein has a certain relationship to proteins which perform completely different functions in

the cell. The sequence analysis of these proteins showed that the genes coding for these proteins are likely due to a common precursor gene or to similar precursor genes. Proteins such as the POM121 protein (Hallberg et al., J. Cell Biol. 122, pages 513-522, 1993) belong to this superfamily. It is one of two known nuclear pore membrane proteins in vertebrates. The CLIP-170 protein which binds vesicles and other organelles within the cell to microtubuli (Pierre et al., Cell 70, pages 887-900, 1992) also belongs to this family. The unexpected discovery that genes which perform such different tasks inside the cell belong to a common protein superfamily is extremely surprising and even inconsistent at first sight. However, when the functions of the individual genes are analyzed, two main functions of these proteins can be derived. The CLIP-170 protein binds to microtubuli, the newly isolated T proteins and the POM121 protein are localized in the nuclear core complex. Due to the properties of these proteins, Inventors propose that this protein superfamily be referred to as POMIC protein superfamily. POMIC shall, in this connection, stand for pores and/or microtubuli-binding protein. Based on the isolation and analysis of the T gene, two paralogs of the T gene could be isolated, namely the T2 and T3 genes which are described in more detail above. As regards evolution and function, the family of the T proteins is between the CLIP (cytoplasmic linker protein-170) and the POM121 protein. This intermediate position is also supported by the sequence analysis and the putative protein structure. The nuclear pore membrane protein POM121 has no marked coiled-coil structure whereas the CLIP-170 protein shows a very distinct coiled-coil structure between the N-terminus and C-terminus (*cf.* figure 29). Coiled-coil structures exist in the family of T proteins, however, they are clearly less marked than in CLIP-170. A similar intermediate position is adopted by the

family of T proteins with respect to the occurrence of hydrophobic domains. The POM121 protein has a hydrophobic domain at the N-terminus which is introduced into the nuclear membrane, and the protein is positioned in the nuclear pore. The CLIP-170 protein has no distinct hydrophobic domain. The T protein and the T3 protein, however, have a hydrophobic domain with three hydrophobic partial regions (*cf.* figure 30). The exchange of the N-terminus in the T2 protein as compared to the evolutionary basic form resulted in a loss of this distinct hydrophobic domain. Yet all three T proteins have in common the very similar structure of the C-terminus. The T3 protein is most similar to the T protein within the T protein family. However, the T3 protein also has undergone a change in the course of evolution. The N-terminus was changed as compared to the T protein by insertion of about 400 amino acids. This insertion resulted in another coiled-coil structure as compared to the otherwise very similar T protein. The T protein and the T3 protein perform functions in the nuclear membrane-localized form, which are similar to those of POM121. However, it is interesting that in the course of evolution there was a loss of part of the C-terminus in the POM121 protein. As compared to the POM121 protein, the T proteins have a longer C-terminus. Due to this longer C-terminus many interactions with other proteins are possible. In this connection, it is also worth mentioning that a leucine-zipper structure was discovered in the T protein, which facilitates interactions with other proteins. The family of T protein plays an important role in the mediation of interactions between cell organelles and filaments, *inter alia* microtubuli. Microtubuli play an important role e.g. in nerve cells; in the case of axons, for example, the plus ends of the microtubuli face away from the cell body whereas the microtubuli of dendrites have both orientations. This

cell polarity is of major importance for the functioning of a cell or living being. Microtubuli also provide an efficient organelle transport, and they are of essential significance for the general organization of membrane structures in a cell. The T proteins perform an important mediator function between membrane structures and microtubuli. The T gene and the T3 gene perform their function in particular as a membrane protein in the nuclear pore whereas the T2 protein acts particularly as a cytoplasmic protein.

Due to the findings of Inventors the T gene and the T3 gene are part of the nuclear pore complex. Nuclear pore complexes (NPCs) are extremely complicated structures which mediate the bi-directional transport of macromolecules between the nucleus and the cytoplasm. The nuclear pore complex is embedded in the nuclear envelope and encases a central channel with a structure only defined insufficiently thus far. Peripheral structures, short cytoplasmic filaments and a basket-like structure are attached on both sides of the central nuclear pore complex. This basket-like structure interacts with molecules which pass through the nuclear pore complex. The mechanism of synthesizing nuclear pore complexes is hardly understood thus far. In addition, it has been found when observing cells passing through mitosis that the nuclear envelope is dissolved deliberately and their components, including the nuclear pore proteins, are distributed over the mitotic cytoplasm. At the end of mitosis, all these components are used again to form the nuclear envelope of the daughter cells. Due to the detailed analysis of the gene T, Inventors found that the N-terminal half of the T protein is weakly homologous to the pore membrane protein POM121. The homology covers the entire region of the POM121 protein and has an identity of about 18

% on a protein level so that the DNAs underlying these proteins should not hybridize with one another, even under hardly stringent conditions. As regards the formation and structure of the nuclear pore, the T protein according to the invention plays a very fundamental role. In a detailed analysis of the protein, a lipophilic domain could be detected at the N-terminus of the T protein. However, this sequence has no homology to the lipophilic sequence of the POM121 protein. There is also a short segment of amino acids which might serve as a signal sequence before the lipophilic domain in the T protein. In order to find out whether this putative signal sequence and the lipophilic domain are involved *in vivo* in the incorporation into the nuclear membrane, various constructs of the T gene were produced. Various parts of the N-terminus of the T protein were fused with the EnhancedGreenFluorescentProtein (EGFP). The EGFP was here fused to the C-terminus of the T protein. The fusion protein which comprised the unchanged N-terminus of the T protein (putative signal sequence with lipophilic membrane domain) was actually incorporated into the nuclear membrane. However, the fusion construct from which the putative signal sequence and the lipophilic domains lack, was not incorporated into the nuclear membrane and accumulated in the cytoplasm. This showed that the N-terminus of the T protein is necessary and suffices to result in a localization within the nuclear membrane. In order to show that the T protein is actually localized in the nuclear membrane, antibodies were generated against a peptide sequence of the T protein. Immunohistochemical studies of tissues of man, mouse and rat were carried out with these antibodies. It showed that the antibody detects a protein which is localized in the nuclear membrane. Since it is difficult to differentiate by means of a light microscope whether the protein is localized in the nuclear membrane or

the nucleus itself, further analyses were made using the high-resolution method of electron microscopy. By this it was possible to clearly show that the T protein is localized in the nuclear membrane. As a detection reaction a second antibody was used here to which the enzyme horseradish peroxidase was coupled and which resulted in a color reaction (DAB). The stain or coloring formed can be seen in the electron-microscopic pictures only on the cytoplasmic side of the nuclear membrane. This indicates that the antibody recognizes an epitope of the T protein which is accessible from the cytoplasmic side for the antibody. The analysis of the immunohistochemical sections also showed that the antibody recognizes very specific neurons (*cf.* figure 24). The results of the analysis of the expression on a protein level by means of the antibody are highly consistent with the results of the analysis of the RNA expression. The mouse ortholog of the T gene was used in the RNA *in situ* analyses. Using the human T gene cDNA clones, murine cDNA clones of the mouse ortholog were initially isolated and sequenced for this purpose. The sequence analysis confirmed that the isolated cDNA clones was the mouse ortholog. Such a murine cDNA clone of the T gene was then used for the RNA *in situ* hybridization (*cf.* figures 25, 26, 27, 28). An expression analysis of the T gene of the mouse was then possible by means of this technique. The accurate analysis of the spatial-temporal expression profile showed that the T gene plays a decisive role in the generation, formation and maintenance of the nervous system in vertebrates. No expression can be detected during the early mouse embryogenesis on day 9.5 post conceptionem (pc = *post conceptionem*). On day 10.5 pc, it is possible to detect an expression in the ventral mesencephalon and in the telencephalon. In this stage there is also a strong expression in the connecting region of the mesencephalon and

telencephalon (forebrain-midbrain). An expression of the T gene in the telencephalon, in the ventral mesencephalon and in the myelencephalon can be detected on day 11.5 pc. An expression in neurons of the mantle zone of the developing brain and in the nuclei of the peripheral nerves is visible on day 12.5 pc. Furthermore, there is an expression in the myelencephalon, spinal cord and spinal ganglia. A minor expression is detectable in the mesencephalon and telencephalon. No expression is detectable e.g. in proliferating neurons in the subventricular layer or in the migrating neurons of the 'intermediate' zone. On day 14.5 pc, an expression in mesenchymal tissues, e.g. around the vertebra or in the region of developing bones, is also visible. A strong expression in all parts of the brain and the peripheral nervous system (e.g. spinal ganglia and nerve fibers of the tail) can be detected on day 16.5 pc. An expression in differentiating neurons of the mantle zone of the telecephalons can also be detected. Furthermore, an expression in neurons of the spinal cord and the spinal ganglia can be detected. When the brain develops after the birth, an expression in the olfactory bulb, in the cerebral cortex and in the developing hippocampus can be detected above all. A minor expression is found however in the coliculus and the developing cerebellum. A similar expression pattern exists in the fully developed brain.

Northern blots (cf. figure 23) were carried out to find out where the T gene or T2 or T3 gene are expressed. The T gene is expressed predominantly in the brain, hardly or not at all in the heart, lungs, placenta, liver, skeletal muscle, kidney or pancreas (irrespective of adult or fetal tissue). However, the T2 gene is virtually not expressed in the brain but strongly expressed in the heart (adult and fetal), adult liver, adult skeletal muscle and adult kidney. The T3 gene

is expressed in all tested tissues (adult and fetal heart, brain, liver, kidney: placenta, adult skeletal muscle, adult pancreas), except in fetal lungs.

Because of the discovery of the T gene and the detailed analysis of this gene with the information obtained therefrom a basis has been created for the development of fully novel medicaments and medicament compound classes. The bi-directional transport of molecules through the nuclear membrane is of decisive significance for the function of each eukaryotic cell. The information which is stored in the form of DNA (chromosomes) in the nucleus is transcribed into mRNA. However, the information is only translated into protein in the cytoplasm. If the transcribed information (mRNA) does not reach the cytoplasm, the information will be lost and dramatic disturbances may occur within the cell. This transport is, however, no one-way street. It is likewise important that certain substances and proteins reach the nucleus so as to maintain the function of the cell. If a transcription factor, for example, which - like the other proteins - is formed in the cytoplasm does not reach the cell nucleus, it cannot trigger the transcription of the other genes. Dramatic disturbances of the events in the cell, which may even comprise the dying of the cell or the organism, are often accompanied by this. This shows clearly that nuclear pore proteins perform an extremely important function within the cell. The analysis of the T gene has now shown that the T protein is also incorporated into the nuclear membrane. It is interesting that the T protein is almost twice as large as the POM121 protein, i.e. it has a much greater binding capacity than the POM121 protein. The T protein is therefore very well suited to isolate possible binding partners which attach to the T protein, in particular to the C-terminus of the T protein.

The tissue-specific expression of the T gene shows strikingly that nuclear core proteins (in particular nuclear pore membrane proteins) do not have to be expressed in all cells and at all times like 'housekeeping' genes. The predominant expression of the T gene in the nervous system shows that the T protein in the nervous system performs a very specific function. The predominant expression of the T gene in the nervous system can now be used for the development of new medicaments and new medicament compound classes. New substances can now be isolated by means of the T protein, which influence deliberately the bi-directional transport in nuclear pores of the nervous system. The localization of the T protein within the nuclear membrane is in this case of major advantage. Chemical compounds can be tested by means of automated tests. Many pharmaceutical companies have suitable screening methods in which more than 200,000 chemicals can be tested. For this purpose, e.g. reporter assays (e.g. GFP fusion proteins, colored substances, etc.) can be used which show the successful transport of a molecule into the nucleus or into the cytoplasm. By this, new active substances can then be isolated which deliberately influence the transport of molecules into nuclear pores, in particular those of the nervous system.

Identifying and analyzing interactions between the T proteins according to the invention (T, T2, T3 protein) or peptides or fragments thereof and possible binding partners which may represent active substances within the above-mentioned meaning, can happen e.g. with the "yeast-two-hybrid system" (Fields, Nature 340, pages 245-247, (1989)). This system is based on the discovery that cellular transcription activators, such as GAL4 or lexA from yeast,

can be separated into two independent functional domains. Both domains are usually part of a protein in the cell nucleus of the yeast cell, which binds to certain activating sequences of different target genes and regulates the transcription thereof. In this connection, one domain, the DNA binding domain (BD), binds specifically to a certain DNA target sequence (upstream activating sequence) in the vicinity of the target promoter. The other domain, the activation domain (AD), increases the transcription rate of the target gene by interaction with the transcription initiation complex which is bound to the promoter of the target gene. In the "yeast-two-hybrid system", this structure is used by the transcription factors in modified form. The DNA binding domain (BD) of GAL4 or lexA is expressed there as fusion protein with a "bait protein or peptide" (here: T, T2 or T3 protein/peptide) in yeast cells. This fusion also has a nuclear localization signal by which it is transported into the cell nucleus of the yeast. The bait fusion protein binds therein to a target sequence (UAS) which is located in the employed yeast strain in the vicinity of the promoters of two reporter genes (e.g. auxotrophic marker (HIS3) and enzymatic marker (lacZ)). By this a constellation results in which the bait protein or peptide is exposed in direct spatial vicinity of the reporter gene promoter. Then, a second fusion protein is additionally expressed in the same yeast cell. It consists of the activation domain (AD) of GAL4 or lexA and a prey protein or peptide. It also has a nuclear localization signal. The prey fusion protein is thus also transported into the cell nucleus of the yeast. If the prey protein and the bait protein exposed on the UAS physically interact with each other, it becomes more likely statistically that the activation domain is located in the vicinity of the reporter gene promoter. This results in an increase of the

transcription of the reporter genes whose extent is proportional to the strength of interaction between bait and prey protein. In this case, e.g. a cDNA library and also a combinatorial peptide library are in consideration as the prey proteins.

The present invention also relates to a process of identifying inhibitors or enhancers of the T protein family according to the invention. For this purpose, the nucleic acid sequences or parts of these sequences, which are part of the T gene or the paralogs or orthologs thereof, are inserted in suitable vectors and used for transfecting or transforming cells, tissues or organisms. These changed cells, tissues or organisms are then used for identifying inhibitors or enhancers of the T protein or its paralog or ortholog proteins (e.g. T2 and T3) or proteins which interact directly or indirectly with these proteins. The inhibitors or enhancers identified by this approach can be used for pharmaceutical active substances or medicaments or for the production thereof and for the treatment of diseases such as cancer, neurological and psychiatric diseases and injuries of the nervous system. In the case of injuries of the nervous system, innate damage of the nervous system or the degenerative diseases of the nervous system, it is possible to support deliberately by this treatment *inter alia* the neuronal regeneration or improve the interconnection of individual nervous regions (used for *inter alia* Alzheimer's disease, Parkinson's disease, schizophrenia, manic-depressive diseases, autism, mental retardation). The present invention provides the possibility of testing the substances or therapeutic agents suitable to enhance or reduce the effect of the T protein or the family of the T proteins. In particular, the changed nuclear pore properties which are influenced by the proteins T and T3 can

be detected by suitable screening methods. The latter include e.g. visualization of the bi-directional transport through the nuclear pore or the detection of a modified transcription of cellular or reporter genes. Substances or therapeutic agents can also be identified which inhibit or promote the effect of proteins which are directly or indirectly involved in the effect of the T protein or the family of the T proteins. Substances or therapeutic agents which show an enhancement or reduction of the effect of the T protein (or T2 or T3) in the above-mentioned screening methods, can be used to determine whether the enhancement or the reduction of the effect of the T protein results in therapeutically desired effects. Above all the inhibition of the growth or spreading of tumor cells or the support of neuronal regeneration, e.g. after injuries of the nerves (*inter alia* paraplegia and head-brain trauma), are counted thereamong. The identified substances can then be used as medicaments or for the production of these medicaments. Due to these medicaments it is then possible to inhibit or block spreading of the disease-inducing cells and thus control or clear up the disease on the whole. An important application of these medicaments is *inter alia* preventing the growth and spreading of tumor cells. In addition thereto, the identified active substances are used as medicaments which stimulate deliberately the growth of certain cells. By this it is then possible to regenerate cells or structures of the nervous system damaged by injury or degenerative processes. The T protein (or T2 or T3) can also be used in screening methods allowing not only to detect the changed nuclear pore properties but also to identify prior or subsequent or parallel signal cascades. By this it is possible to identify e.g. tyrosine kinases or tyrosine phosphatases which regulate proteins which in turn influence directly or indirectly the action of the T protein (or T2 or T3). As a

result, suitable targets for the positive influence of the events in the cells can be recognized and characterized. Furthermore, the T protein, although it occurs as a nuclear pore protein, is significant for the interactions with filaments of the cell, e.g. microtubuli and actin. These interactions can now be studied, e.g. by fusion proteins of the T protein with the EGFP protein. Cells which were stably or transiently transformed or transfected with constructs for such fusion-reporter proteins, can be incubated with substances or pharmaceutical preparations to identify substances which enhance or reduce the interaction of the T protein with filaments such as the actin filaments or the microtubuli. As a result, it is possible to isolate active substances which positively influence *inter alia* the growth of nerve cells or the inhibition of the growth of tumor cells. For example, immunoprecipitation has to be mentioned as a method of identifying such possible active substances. Proteins can be isolated by this which bind to the T protein family. Further immunoprecipitations can then be carried out with these proteins to isolate new proteins which then no longer interact directly with the T protein.

The present invention also relates to a method of identifying further proteins which play a role in the development and function of the nervous system and/or are a nuclear pore protein, the method comprising the steps of:

- (a) producing an antibody against a protein of the T family (T, T2 or T3 protein),
- (b) contacting a cell extract with the antibody and identifying the antibody/protein complex,

- (c) analyzing the complex to identify a protein which has bound to the protein of the complex and is no antibody, and
- (d) optionally repeating steps (a) to (c) to identify further proteins of this function.

The invention is described in more detail by means of the figures, which show:

Figure 1: human cDNA sequence (gene T) and derived amino acid sequence

Figure 2: human genomic DNA sequence (gene T)

Figure 3: human genomic DNA sequence (gene T)

Figure 4: human genomic DNA sequence (gene T)

Figure 5: human genomic DNA sequence (gene T)

Figure 6: human genomic DNA sequence (gene T)

Figure 7: human genomic DNA sequence (gene T)

Figure 8: human genomic DNA sequence (gene T)

Figure 9: partial murine cDNA sequence (gene T) and derived amino acid sequence

Figure 10: partial murine genomic DNA sequence (gene T)

Figure 11: partial human cDNA sequence (gene T2) and derived amino acid sequence

Figure 12: partial murine cDNA sequence (gene T2) and derived amino acid sequence

Figure 13: partial murine cDNA sequence (gene T2) and derived amino acid sequence

Figure 14: splicing variant of the human T gene with derived amino acid sequence

Figure 15: splicing variant of the human T gene with derived amino acid sequence

Figure 16: partial human cDNA sequence (gene T2) with derived amino acid sequence

Figure 17: partial human cDNA sequence (gene T3; protein isoform 1) with derived amino acid sequence

Figure 18: partial human cDNA sequence (gene T3; protein isoform 2) with derived amino acid sequence

Figure 19: partial murine cDNA sequence with derived amino acid sequence (gene T3)

Figure 20: oligonucleotide and peptides (T gene)

Figure 21: sequence comparison within the T family

Figure 22: protein alignment of POM121 protein and T protein

Figure 23: Northern blot analysis

Figure 24: immunohistochemical studies and electron-microscopic pictures

Figure 25: *in situ* hybridization with embryonal RNA

Figure 26: *in situ* hybridization with RNA from brain

Figure 27: *in situ* hybridization with RNA from fetal brain

Figure 28: *in situ* hybridization with RNA from nerve tissues of mouse

Figure 29: comparison of the coiled-coil regions between CLIP protein, T protein and POM121

Figure 30: hydrophobicity blot for POM121, T protein and T3 protein.

The following clones were deposited with the DSMZ (*Deutsche Sammlung für Mikroorganismen und Zellkulturen GmbH*) [German-type collection of microorganisms and cell cultures], Mascheroder Weg 1b, Braunschweig, according to the Budapest treaty on August 18, 1998:

- clone JFC277 (DSM12371); human cDNA; represents the human cDNA sequence of Bp 1218-3690
- clone JFC405 (DSM12372); human cDNA; represents the human cDNA sequence of Bp 1-1891
- clone JFC601 (DSM12373); murine cDNA; represents the murine cDNA sequence of Bp 225-3026
- clone JFC950 (DSM12374); human genomic clone; represents human genomic sequence

- clone JFC955 (DSM12375); human genomic clone; represents human genomic sequence; comprises start of the cDNA sequence
- clone JFC N2112 (DSM12376); human genomic clone; was fully sequenced. The sequence is shown in figure 2 and contains the sequence of Bp 1756-4228 of the human cDNA sequence.

The following clone was deposited with DSMZ according to the Budapest treaty on February 2, 1999:

- clone JFC-BN27 (DSM 12659); contains the sequence of Bp 4370-8690 of the human cDNA sequence.

The following clone was deposited with the DSMZ according to the Budapest treaty on February 19, 1999:

- clone JFC-BN20 (DSM 12698); contains the sequence of Bp 2025-6280 of the human cDNA sequence

The following clone was deposited with the DSMZ according to the Budapest treaty on February 1, 2000.

- cDNA clone pL70 (DSM13270); represents essential parts of the gene T3.

The sequences shown in figures 2 to 8 originates from clones JFC955 (DSM 12375) and JFC950 (DSM 12374). The sequence shown in figure 1 originates from clones JFC277 (DSM 12371), JFC405 (DSM 12372) and JFC-BN27 (DSM 12659) and JFC-BN20 (DSM 12698). The sequence shown in figure 9 originates from the clone JFC610 (DSM12373).

The invention is further described by means of the following embodiment.

### **EXAMPLES**

As to the methods employed reference is also made to Sambrook, J., Fritsch, E.F. and Maniatis, T. (Molecular Cloning; A Laboratory Manual; second edition; Cold Spring Harbor Laboratory Press, 1989) and Current Protocols in Molecular Biology (John Wiley and Sons, 1994-1998), the below techniques, in particular preparation of DNA or RNA or Northern blot, being sufficiently known to, and mastered by, the person skilled in the art.

Before it is described in detail how the experiments are carried out, the operating strategy is to be explained first.

When screening for genes triggering diseases of the CNS (e.g. neurodegenerative diseases, mental retardations, tumoral diseases of the CNS) in the mutated state, 23 cDNA clones were isolated from a human fetal brain cDNA library (Stratagene company, Heidelberg). A human fetal brain cDNA library was used as a starting material, since it was assumed that genes which play a role in the development of the CNS and in particular of the brain are present in a fetal brain cDNA library. However, since what is called housekeeping genes (genes expressed in most tissues) are also expressed in the CNS, it was tested whether the select cDNA clones originate from genes having a CNS-specific expression. For this purpose, the cDNA pieces ('inserts') contained in the individual cDNA clones were isolated and used for hybridization with Northern blots. The employed

Northern blots comprised polyA RNA from different human tissues (e.g. brain, skeletal muscle, liver and kidney) and various development stages (fetal and adult tissues). Since as mentioned above not only brain-specific genes are expressed in the brain, the hybridization with the Northern blots was used to identify cDNA clones which are expressed above all in the brain and not so much in other tissues. Due to this differential analysis it was possible to identify a cDNA clone which has a brain-specific expression pattern. Using this cDNA clone, the entire mRNA sequence for the new protein encoded therein could be isolated and deciphered (gene T with protein TP encoded therein) by repeated hybridization of the fetal cDNA library.

**EXAMPLE 1: Identification of the T genes**

**1. Titration of the cDNA libraries**

In order to ensure an effective infection, it was initially necessary to produce phage-competent bacteria in an overnight culture. The magnesium ions contained in the medium induce the maltose receptor of the bacteria to which the phage binds to infect the bacterium.

Performance:

Charge 50 µl *E. coli* XL1-Blue in 50 ml LB broth, the medium being admixed with MgSO<sub>4</sub> in a concentration of 10 mM. Incubate overnight at 30°C and 220 rpm. Centrifuge off the bacteria at 4°C and 1000 xg for 10 min. Resuspend in 25 ml 10 mM MgSO<sub>4</sub>. The thus produced phage-component bacteria could be stored at 4°C for up to one week.

**2. Culturing the cDNA libraries**

For culturing the library, Baltimore Biological Lab. (BBL) agar plates and BBL top agarose had to be prepared. The phages (human or murine cDNA library, Stratagene company) were mixed with SM medium to a dilution of 1:10<sup>3</sup> and 1:10<sup>4</sup> to obtain individual plaques after the culturing.

Performance:

For the BBL agar (pH 7.2) 10 g BBL trypticase, 5 g NaCl and 10 g Select agar were weighed and filled to 1 l with H<sub>2</sub>O. The agar is dissolved by autoclaving. After cooling to about 60° pour the plates. The plates are preheated to 37°C prior to their use to avoid premature solidification of the top agarose. The BBL top agarose (pH 7.2) was prepared with 10 g BBL trypticase, 5 g NaCl, 6.5 g agarose and 10 ml 1 M MgSO<sub>4</sub> solution. Dissolve by autoclaving and provide in the water bath to 41°C. Add 15 µl of the above indicated dilute phage solution and 250 µl of the competent XL-1 bacteria in a 15 ml Falcon tube. Incubate at room temperature for 20 minutes. Add 10 ml BBL top agarose, swivel and place on the heated agar plate. The top agarose layer is solid after about 20 minutes and the plates can be stacked with the agar side up. Incubation is carried out overnight at 37°C. The plates can be stored at 4°C after expired incubation time or can be used directly for transferring the phage plaques. Carefully close the plates for storing them together with a chloroform-soaked cloth in plastic bags. The chloroform prevents the growth of cryophilic bacteria and fungi.

3. In vivo excision

The employed cDNA libraries (human and murine fetal brain cDNA library; Stratagene company, Heidelberg) were cloned in the vector λ-ZAPII. Due to this there was the possibility of circumventing the subcloning of the phage insert in a plasmid vector. This protocol permits to transfer cDNA which

is located as insert in the  $\lambda$ -ZAPII vector into an insert in simple way by an *in vivo* preparation which is now found in the plasmid Bluescript SK(-). In principle, this preparation serves for introducing by a helper phage information for proteins which permit DNA amplification only in the region of the phage genome, which have the genetic information for the plasmid with cDNA insert. For the most part, the method was carried out in accordance with the protocol of the manufacturer (Stratagene).

In particular, culturing was made such that individual phage plaques were on the plate. Then, the *in vivo* excision protocol was carried out with these individual plaques. The plasmid DNA and its plasmid inserts were isolated from the bacterial clones and subsequently hybridized with Northern blots. The selection of further clones to be studied was based on the expression pattern in the Northern blots.

#### Performance:

Mix 100  $\mu$ l of a single phage  $\lambda$ -ZAPII clone with 200  $\mu$ l XL1 bacteria and 2  $\mu$ l helper phages (contained in the Stratagene kit). Shake for 15 min. at 37°C and 80 rpm, the specific attachment of both phage types to the host bacterium taking place. Add 3 ml LB broth. Incubate for 2 h at 37°C and 200 rpm. The DNA replication of the plasmid contained in the  $\lambda$ -ZAPII, its circularization and the packing into coat proteins take place and discharge from the bacterium occur during this time. Heat to 70°C for 20 minutes. Thereafter, centrifuge at 4000 g for 15 minutes. This kills the still remaining bacteria and separates their fragments from the plasmids existing in the phage coat, which are found in the supernatant. Add 1  $\mu$ l thereof to 200  $\mu$ l SOLR host cells, incubate at 37°C for 15 minutes. Plate 100  $\mu$ l onto LB/Amp plates. Store at 37°C overnight. The then grown bacterial

clones contain the plasmid with the corresponding cDNA insert. A mini-prep DNA preparation was carried out each.

#### **4. "random primed" DNA labeling**

The radioactive labeling of the double-stranded insert DNA of the cDNA clone was carried out as follows for the further isolation of overlapping cDNA clones:

Performance:

Dissolve 100 ng DNA in a volume of 12  $\mu$ l H<sub>2</sub>O for a typical labeling batch. 10-minute heating to 95°C effects the denaturation of the DNA into single strands. Store the preparation on ice to prevent reassociation of the two complementary DNA strands. Complete the reaction batch by 4  $\mu$ l OLG (oligo-labelling buffer), 1  $\mu$ l Klenow (1U) and 2.5  $\mu$ l  $\alpha$ -<sup>32</sup>P-dCTP and 2.5  $\mu$ l  $\alpha$ -<sup>32</sup>P-dATP. Incubate at room temperature overnight. Based on the hexanucleotides attached to a single strand, the formation of the complementary strand takes place during this time by the Klenow fragment of the *E. coli* DNA polymerase I. The DNA is labeled radioactively by incorporating  $\alpha$ -<sup>32</sup>P-dCTP and the  $\alpha$ -<sup>32</sup>P-dATP.

#### **5. Separation of non-incorporated radioactive nucleotides**

The non-incorporated nucleotides were separated by means of a personally prepared sephadex G-50 column. The separation principle of the column is based on the exclusion chromatography. The smaller non-incorporated nucleotides fit into small pores of the column material while the DNA is locked out. The volume in which the nucleotides may move is thus greater than the volume available to the DNA. If a mixture of DNA and nucleotides is placed on the column, the DNA runs through the column faster than the nucleotides. This permits the separation of non-incorporated nucleotides.

**Performance:**

A Pasteur pipette was closed with a small glass bead. Fill the Pasteur pipette with sephadex G-50 ("fine") dissolved in water until the filling material is 5 cm below the top edge of the Pasteur pipette. Rinse the column 2 times with TE. Apply the above radioactive labeling batch. Add 320  $\mu$ l TE. Discard the solution which has run through the column. Place an Eppendorf tube below the column. Add 350  $\mu$ l TE. Collect the radioactive solution run through the column.

**6. Plaque "blot"**

The plaque "blot" was made to analyze the cDNA library to make accessible the cDNA in the phage clones to hybridization.

**Performance:**

Place a labeled hybond-N membrane provided with an inscription in air bubble-free manner on the plate with the phage plaques for one minute. The labeling pattern was transferred. Place it on a Whatman paper soaked with denaturing solution (0.5 M NaOH; 1.5 M NaCl) for 10 minutes. Neutralize in 50 mM phosphate buffer for 10 minutes. The rests of the bacterial layer are wiped off with slight pressure using a phosphate buffer-soaked Kleenex cloth. The filters are spread at room temperature for drying. Thereafter, the filters were baked at 90°C for 1 h.

**7. Hybridization**

The hybridization is based on the binding of complementary, single-stranded nucleic acids. For this purpose, the DNA to be studied was immobilized on a membrane and hybridized with a radioactively labeled probe. The complementary binding is maintained even after washing off the non-specifically adhering probes and can be made visible by means of

autoradiography. Single-stranded molecules were incubated during the hybridization under salt and temperature conditions which support the formation of base-paired double strands. A decisive factor in the association and dissociation kinetics are the hydrogen bridge bonds between the base pairs G-C and A-T. The hybridization reaction is influenced by changing the temperature and the salt and sample concentrations.

Performance:

First, prehybridize the DNA filters in hybridization solution (0.5 M NaPi (pH 7.2); 7 % SDS; 0.2 % BSA; 0.2 % PETG 6000; 0.05 % polyvinyl pyrrolidone 360000; 0.05 % Ficoll 70000; 0.5 % dextrane sulfate) with a 0.1 ml/cm<sup>2</sup> at 65°C. For this purpose, incubate the filters in a plastics box in a shaking water bath at 65°C for a period of at least 1 h. Discard the prehybridization solution. Place the radioactively labeled sample (see above items 4. and 5.) with 0.5 ml/cm<sup>2</sup> of hybridization solution (65°C) on the filters. The activity of the sample should not drop below 50 cpm, measured at a distance of 40 cm. The hybridization takes place overnight at 65°C (human cDNA library) or 55°C (interspecies hybridizations man-mouse and for isolating the homologous genes). Wash the filters two times for 30 minutes with about 500 ml wash buffer in a shaking bath at 65°C (55°C). Autoradiography was then carried out.

**8. Autoradiography**

The filters were packed in plastic foodwrap. The autoradiography was made at -80°C in an X-ray cassette containing a reinforcing film made of calcium tungstate. The exposure is 30 minutes to several days, depending on the strength of the signal.

The complete mRNA which codes for the protein of the T gene could be isolated by means of the above mentioned techniques. Furthermore, using cDNA clones of this newly isolated T gene it was possible to isolate two further genes (T2 and T3) which have distinct homologies with this gene. For this purpose, the above mentioned techniques were used again. For isolating the related genes T2 and T3, the hybridization temperature was lowered to 55°C.

#### **EXAMPLE 2: Northern blot**

The 'multiple tissue Northern blots' were purchased from the CLONTECH company (Palo Alto, California, U.S.A.) and used in accordance with the instructions from the manufacturer. The respective DNA samples of the genes T, T2 and T3 were labeled radioactively and hybridized with the Northern blots. The sequence of bp 1-4200 of figure 1 was used for the analysis of the expression pattern on a Northern blot level. For the gene T3 the sequence of bp 1310-4870 of figure 17 was used for hybridization. The sequence of bp 3120-4230 of figure 16 was used for the gene T2. The "random priming" method was used for the radioactive labeling of double-stranded DNA.

##### a) Random priming:

Dissolve 100 ng DNA in a volume of 12 µl for a typical labeling batch. 10-minute heating to 95°C effects the denaturation of the DNA into single strands. Store the batch on ice to prevent reassociation of the two complementary DNA strands. Complete the reaction batch by 4 µl OLB, 1 µl Klenow (1U) and 2.5 µl  $\alpha$ -<sup>32</sup>P-dCTP and 2.5 µl  $\alpha$ -<sup>32</sup>P-dATP. Incubate at room temperature overnight. Based on the hexanucleotides attached to a single strand, the formation of the complementary strands takes place during this time by

the Klenow fragment of the *E. coli* DN polymerase I. The DNA is labeled radioactively by the incorporation of the  $\alpha$ -<sup>32</sup>P-dCTP and the  $\alpha$ -<sup>32</sup>P-dATP.

The non-incorporated nucleotides were separated by means of a personally prepared sephadex G-50 column. The separation principle of the column is based on the exclusion chromatography. The smaller non-incorporated nucleotides fit into small pores of the column material while the DNA is locked out. The volume in which the nucleotide may move is thus greater than the volume available to the DNA. If a mixture of DNA and nucleotides is placed on the column, the DNA runs through the column faster than the nucleotides. This permits the separation of non-incorporated nucleotides. For this purpose, a Pasteur pipette is closed with a small glass bead. Fill the Pasteur pipette with sephadex G-50 ("fine") dissolved in water until the filling material is 5 cm below the top edge of the Pasteur pipette. Rinse the column 2 times with TE. Apply the above radioactive labeling batch. Add 320  $\mu$ l TE. Discard the solution which has run through the column. Place Eppendorf tube below the column. Add 350  $\mu$ l TE. Collect the radioactive solution run through the column.

b) Hybridization:

The Northern blots were hybridized as described below. First, the Northern blots were prehybridized at 65°C in 10 ml hybridization solution (350 ml 20 % SDS, 500 ml 1 M phosphate buffer, pH 7.2; 150 ml distilled water). For this purpose, the Northern blots were incubated in a glass tube in a hybridization roll-over-type furnace at 65°C for a period of 6 h.

The prehybridization solution was discarded. The radioactively labeled sample was placed with 10 ml hybridization solution (65°C) on the filters.

The hybridization was carried out at 65°C overnight. The filters were then washed two times for 30 min. with about 500 ml wash buffer (80 ml 1 M phosphate buffer, pH 7.2; 100 ml 20 % SDS, 1820 ml distilled water) at 65°C in a shaking bath.

c) Autoradiography

The filters were welded into plastic film. The autoradiography was made at -80°C in an X-ray cassette which contained a reinforcing film of calcium tungstate. Exposure was 1 to 4 days depending on the strength of the signal.

The results of the Northern blots carried out are shown in figure 23.

**EXAMPLE 3: RNA *in situ* hybridization**

Embryos in various development stages were isolated from pregnant NMRI mice. The embryos and other tissue samples were fixed overnight with 4 % paraformaldehyde in PBS at 4°C. 10 µm freezing sections of the embryos were transferred to slides coated with 3-aminopropyl triethoxysilane. Sense strand ("sense") and antisense strand ("antisense") samples were produced by transcription with  $\alpha^{35}\text{S}$ -UTP with a specific activity of  $>10^9$  decays per minute/µg. For this purpose, the linearized mouse T gene cDNA clone from figure 9 was transcribed with T7 or Sp6-RNA polymerase. The sample length was reduced by alkaline lysis to 150 to 200 nucleotides. The slides were prehybridized at 54°C in a solution containing 50 % formamide, 10 % dextrane sulfate, 0.3 M NaCl, 10 mM Tris, 10 mM sodium phosphate, pH 6.8, 20

mM dithiothreitol, 0.2 % Denhardt's solution, 0.1 Triton X-100, 0.1 mg/ml Escherichia coli RNA and 0.1 mM non-radioactive  $\alpha$ -S-UTP. The  $^{35}$ S-labeled sample ( $8 \times 10^4$  decays per minute per ml) were added to the hybridizing mixture for the hybridization and the hybridization was then continued for 16 h at 54°C in a humid chamber. The slides were then washed in the hybridization solution for 2 hours. The remaining non-hybridized RNA sample was then digested using RNase A. Thereafter, the slides were washed for 30 minutes at 37°C with 2x SSC, 0.1 % SDS and for 30 minutes with 0.1x SSC, 0.1 % SDS. Then, the slides were dehydrated with increasing ethanol concentrations. The slides were covered with Ilford K5 autoradiography emulsion. After 1 to 2 weeks of exposure at 4°C, the slides were incubated in Kodak D19b developer and dyed with Giemsa. The sections were analyzed in dark field and bright field illumination with a Zeiss SV8 stereomicroscope and an Axiophot microscope and photographed with an Agfa ortho black-and-white film.

The results of the RNA *in situ* hybridization are shown in figures 25, 26, 27 and 28.

Figure 25: expression of the murine T gene during the mouse embryogenesis. Bright field (a,c,e,g) and dark field pictures (b, d, f, h) of horizontal (a,b) and sagittal sections (c-h) through a 10.5 (a,b), 12.5 (c,d), 14.5 (e,f) and 16.5 (g,h) dpc embryo (dpc = days post conceptionem) which were hybridized with an antisense ribo sample of the murine T gene. Dec = decidua, g = guts, he = heart, lab = labyrinth, li = liver, me = myelcephalon, sc = spinal cord, sga = spinal ganglia, sb = tooth bud, te = telencephalon. Bar = 1 mm.

Figure 26: Expression of the murine T gene in the postnatal brain. Bright field (a,d) and dark field pictures (b,c,e,f) of horizontal sections through an 1 wpn (weeks *post natalis*) and 6 wpn head, which were hybridized with a T gene antisense (b,e) and a sense sample (c,f). cd = cerebellum, cor = cortex, cos = colliculus, ey = eye, hi = hippocampus, ne = nasal epithelium, ob = olfactory bulb, bar = 1 mm.

Figure 27: Greater enlargement of the 10.5 dcp embryo of figure 25 a,b. The arrows point to a region of little expression in the somites (arrows in b). An intense expression can be seen in the region between mesencephalon and telencephalon ("forebrain-midbrain junction"). Aod = aorta dorsalis, me = mesencephalon, sc = spinal cord, te = telencephalon. Bar = 100  $\mu$ m.

Figure 28: Expression of the T gene during the development of the nervous system. Expression of the T gene in neurons of the mantle zone of the developing brain and in nuclei of peripheral nerves (arrow in b). No expression is visible in proliferating neurons in the subventricular layer or in migrating neurons of the intermediate zone (c,d). On day 16.5, an intense expression is visible in differentiating neurons of the mantle zone of the telencephalon (e,d). A minor expression is also visible in neurons of the spinal cord and the spinal ganglia (g,h). Furthermore, a minor expression is visible in an individual layer below the skin (g,h). iz = intermediate zone, mz = mantle zone, sc = spinal cord, sga = spinal ganglia, sk = skin, svl = subventricular layer, vn = ventricle. Bar = 100  $\mu$ m.

**EXAMPLE 4: Production of antibodies**

Using a synthetically produced peptide of the sequence "EKGEDPETRRMRTVKNIAD" animals are immunized to produce antibodies against the T protein as follows:

Immunization protocol for polyclonal antibodies in rabbits

600 µg purified KLH-linked peptide in 0.7 ml PBS and 0.7 complete or incomplete Freund's adjuvant are used per immunization:

Day 0: 1<sup>st</sup> immunization (complete Freund's adjuvant)  
Day 14: 2<sup>nd</sup> immunization (incomplete Freund's adjuvant; icFA)  
Day 28: 3<sup>rd</sup> immunization (icFA)  
Day 56: 4<sup>th</sup> immunization (icFA)  
Day 80: bleeding to death.

The rabbit serum is tested in an immunoblot. For this purpose, the protein used for the immunization is subjected to SDS polyacrylamide gel electrophoresis and transferred to a nitrocellulose filter (*cf.* Khyse-Andersen, J., J. Biochem. Biophys. Meth. 10 (1984), 203-209). The Western blot analysis was carried out as described in Bock, C.-T. et al., Virus Genes 8, (1994), 215-229. For this purpose, the nitrocellulose filter is incubated with a first antibody at 37°C for one hour. This antibody is the rabbit serum (1:10000 in PBS). After several wash steps using PBS, the nitrocellulose filter is incubated with a second antibody. This antibody is an alkaline phosphatase-coupled monoclonal goat anti-rabbit IgG antibody (Dianova company) (1:5000) in PBS. 30 minutes of incubation at 37°C are followed by several wash steps using PBS and subsequently by the alkaline phosphatase detection reaction with developer solution (36 µM 5'-bromo-4-chloro-3-indolylphosphate, 400 µM nitro blue tetrazolium, 100 mM Tris-HCl, pH 9.5, 100 mM

NaCl, 5 mM MgCl<sub>2</sub>) at room temperature until bands become visible.

It shows that polyclonal antibodies according to the invention can be prepared.

Immunization protocol for polyclonal antibodies in chickens

100 µg of purified KLH-linked peptide in 0.8 ml PBS and 0.8 ml of complete or incomplete Freund's adjuvant are used per immunization.

Day 0: 1<sup>st</sup> immunization (complete Freund's adjuvant)

Day 28: 2<sup>nd</sup> immunization (incomplete Freund's adjuvant; icFA)

Day 50: 3<sup>rd</sup> immunization (icFA)

Antibodies are extracted from egg yolk and tested in a Western blot. Polyclonal antibodies according to the invention are detected.

Immunization protocol for monoclonal antibodies in mice

250 µg of purified KLH-coupled peptide in 0.25 ml PBS and 0.25 ml of complete or incomplete Freund's adjuvant are used per immunization. The peptide is dissolved in 0.5 ml (without adjuvant) in the 4<sup>th</sup> immunization.

Day 0: 1<sup>st</sup> immunization (complete Freund's adjuvant)

Day 28: 2<sup>nd</sup> immunization (incomplete Freund's adjuvant; icFA)

Day 56: 3<sup>rd</sup> immunization (icFA)

Day 84: 4<sup>th</sup> immunization (PBS)

Day 87: fusion.

Supernatants of hybridomas are tested in a Western blot. Monoclonal antibodies according to the invention are identified.

**EXAMPLE 5: immunohistochemical studies**

The immunohistochemical studies shown in figure 24 were made with an affinity-purified polyclonal rabbit antibody, produced above, against the T protein (referred to as first antibody below). Mouse brain was removed and treated as follows:

1<sup>st</sup> day

section thickness 6-10 µm, common fixation on slides, storage at -80°C for up to about 2 months

Take out the sections the evening before and allow them to dry at room temperature overnight

Rinse slides in PBS, pour off, rinse once again, thereafter allow to stand in PBS for 10 min.

Take out slides and wipe off the liquid around the tissue using a cloth.

Encircle using PAP-PEN (protein-glycerol; Dako company) so that no more liquid can flow out.

Add 100 µl peroxidase blocking solution (Dako company, Hamburg), incubate for 20 minutes.

Rinse slides in PBS, pour off, rinse again, thereafter allow to stand in PBS for 10 min.

Take out slides and wipe off the liquid around the tissue using a cloth.

Prepare an 1:10 dilution of normal (sheep) serum in PBS (e.g. sheep Dako X0503, Dako company, Hamburg), add 100 µl thereof and incubate for 20 minutes.

Rinse slides in PBS, pour off, rinse again, thereafter allow to stand in PBS for 10 minutes.

Take out slide and wipe off the liquid around the tissue using a cloth.

Add first antibody in a dilution of 1:100.

Add 100 µl of the first antibody (in PBS) and incubate in a refrigerator in a humid chamber overnight. Control: without first antibody.

2<sup>nd</sup> day

Take humid chamber out of the refrigerator and allow to stand at room temperature. Rinse slide in PBS, pour off, rinse again, thereafter allow to stand in PBS for 10 minutes, when many slides are analyzed wash two times with PBS.

Take out slides and wipe off the liquid around the tissue using a cloth.

Prepare a 1:100 dilution of second antibody "antirabbit biotinylated" (Amersham company, Braunschweig) in PBS and add 100 µl thereof.

Incubate in a humid chamber at room temperature for 45 minutes.

Rinse slides in PBS, pour off, rinse again, thereafter allow to stand in PBS for 10 minutes.

Take out slide and wipe off the liquid around the tissue using a cloth.

Prepare a 1:100 dilution of streptavidine peroxidase (streptavidine horseradish) (Amersham company, Braunschweig) with PBS and add 100 µl thereof.

Incubate in a humid chamber at room temperature for 45 minutes.

Rinse slides in PBS, pour off, rinse again, thereafter allow to stand in PBS for 10 minutes.

Take out slides and wipe off the liquid around the tissue using a cloth.

Staining: Add one drop chromogen per ml buffer just before the use. Vortex and place in the dark.

Add 100 µl staining solution (Dako company, Hamburg).

Finally, stain the control. Incubate for about 2 minutes.

Incubate slides in water. Inspect under a microscope.

Place 1-2 drops of crystal Mount on the section. If there is an air bubble, suck it off with a paper handkerchief.

The rest of the slide is wiped off using HCl-EtOH to remove the stain.

Place a line of adhesive (Eukitt) on the cover glass. Press the cover glass onto the slide without producing air bubbles.

The enzyme in the second antibody results in a dye formation (DAB) so that the T protein can be detected.

Figure 24 (a-d): Light-microscopic pictures which show that the T protein is localized in or at the nucleus of the cell. The electron-microscopic picture in e shows that the T protein is not localized in the nucleus but in the membrane. The pictures are highly consistent with a function as a membrane-terminal nuclear pore protein. The arrows in e show the stain formed which can be seen on the cytoplasmic side of the nuclear membrane.

•

**Claims**

1. DNA sequence coding for a protein which is involved in the development of the nervous system, in particular the CNS, and is expressed in a tissue-specific and development-specific manner, wherein the DNA sequence comprises the following DNA sequences:
  - (a) the DNA sequence of figure 1, figure 2, figure 3, figure 4, figure 5, figure 6, figure 7 or figure 8;
  - (b) the DNA sequence of figure 9 or figure 10;
  - (c) the DNA sequence of figure 11;
  - (d) the DNA sequence of figure 12 or figure 13;
  - (e) the DNA sequence of figure 14 or figure 15;
  - (f) the DNA sequence of figure 16;
  - (g) the DNA sequence of figure 17 or 18;
  - (h) the DNA sequence of figure 19;
  - (i) a DNA sequence hybridizing with (a), (b), (c), (d), (e), (f), (g) or (h);
  - (j) fragments, variants, functional equivalents, derivatives or precursors of the DNA sequence of (a), (b), (c), (d), (e), (f), (g), (h) or (i); or
  - (k) a DNA sequence which differs from the DNA sequence of (a), (b), (c), (d), (e), (f), (g), (h), (i) or (j) due to the degeneration of the genetic code.
2. The DNA sequence according to claim 1, which codes for a protein or peptide comprising the amino acid sequence of figure 1, figure 9, figure 11, figure 12, figure 13, figure 14, figure 15, figure 16, figure 17, figure 18 or figure 19, wherein the protein or peptide has the biological activity defined in claim 1.

3. An antisense RNA, characterized in that it is complementary to the DNA sequence of claim 1 or 2 and can reduce or inhibit the synthesis of the protein encoded by this DNA sequence.
4. Ribozyme, characterized in that it is complementary to the DNA sequence of claim 1 or 2 and can bind specifically to the RNA transcribed by this DNA sequence and can cleave it so as to reduce or inhibit the synthesis of the protein encoded by this DNA sequence.
5. Expression vector, containing the DNA sequence according to claim 1 or 2 or coding for the antisense RNA according to claim 3 or the ribozyme according to claim 4.
6. The expression vector according to claim 5, which comprises additionally the promoter of the human T gene or an ortholog of the T gene.
7. Expression vector according to claim 5 or 6, which codes for the T, T2 or T3 proteins or for fragments thereof in the form of a reporter fusion protein.
8. Host cell which is transformed with the expression vector according to any of claims 5 to 7.
9. Protein which is encoded by the DNA sequence according to claim 1 or 2 and which is involved in the development of the nervous system and is expressed in tissue-specific and development-specific manner, or fusion proteins, fragments, variants, derivatives or precursors of the protein.

10. Protein according to claim 9, which has one of the following motives:

Motive 1:

(A, T) (I, P, V) (L, T) (G, A, Q) (L, V) XXX (L, V)

Motive 2:

IYTDQWAN

Motive 3:

XXXXXXXXXXGXXXXXXAXXXXXXXXXXXXXXXXXXXXXQ

Motive 4:

SXXXXDX(12,20)KX(17,22)AXXXXXXXXL

Motive 5:

IYTDWANXXLX(K, R)

Motive 6:

KX(18,21)AXXXXXXXXLX(15,24)S

Motive 7:

NX(3,11)SXXXAXXXXXXXL

wherein X = every amino acid

(A, T) = amino acid A or T at this site

X(number 1, number 2) = number 1 to number 2

Xs at this site

11. Method of producing the protein according to claim 9, which comprises culturing the host cell according to claim 8 under suitable conditions and obtaining the protein from the cell or the culture medium.

12. Antibody which is directed against the protein according to claim 9 or fragment thereof.

13. Antibody according to claim 12, which is obtained by immunizing animals with a peptide having the sequence "EKGEDPETRRMRTVKNIAD".

14. Use of the DNA sequence according to claim 1 or 2, the antisense RNA according to claim 3, the ribozyme according to claim 4, the expression vector according to any of claims 5 to 7, the protein according to claim 9 or the antibody or the fragment thereof according to claim 12 or 13 for preventing or treating diseases of the nervous system, in particular of the CNS.
15. Use according to claim 14, wherein the disease of the nervous system is a tumoral disease of the CNS.
16. Use according to claim 14, wherein the treatment of diseases of the nervous system are the promotion of the neuronal regeneration in the case of injuries of the nervous system and degenerative diseases of the nervous system.
17. Use according to claim 14, wherein the treatment of diseases of the nervous system are the regeneration of the neuronal linkages and the regeneration of the innate and acquired malfunctions of the nervous system.
18. Use according to claim 15 for inhibiting the growth and spreading of tumor cells.
19. Diagnostic method for detecting a disturbed expression of the protein according to claim 9 or for detecting a changed form of this protein, in which a sample is contacted with the DNA sequence according to claim 1 or 2 or the antibody or the fragment thereof according to claim 12 or 13 and then it is determined directly or indirectly whether the concentration of the protein

and/or its amino acid sequence differs from a protein obtained from a healthy patient.

20. Diagnostic kit for carrying out the method according to claim 19, which contains the DNA sequence according to claim 1 or 2 and/or the antibody or the fragment thereof according to claim 12 or 13.
21. Non-human mammal whose naturally occurring T, T2 or T3 gene comprises a change in the gene structure or the gene sequence.
22. Non-human mammal, wherein a change of the gene structure of the T, T2 or G3 gene is achieved in the mammal by introducing a deletion in place of which a homologous or heterologous sequence is introduced.
23. Non-human mammal, wherein a change of the gene structure of the T, T2 or G3 gene is achieved by inserting a homologous or heterologous sequence in the corresponding gene naturally occurring in the mammal.
24. Non-human mammal according to claim 22 or 23, wherein the heterologous sequence is the selection marker sequence.
25. Non-human mammal according to claim 24, wherein the selection marker sequence conveys resistance to neomycin.
26. A method of producing a non-human mammal according to any of claims 21 to 25, characterized by the steps of:
  - (a) producing a DNA fragment, in particular a vector, containing a changed T, T2 or G3 gene, the T, T2

or T3 gene having been modified by inserting a heterologous sequence, in particular a selectable marker;

- (b) preparing embryonal stem cells from a non-human mammal (preferably a mouse);
- (c) transforming the embryonal stem cells from step (b) with the DNA fragment from step (a), the T gene in the embryonal stem cells being changed by homologous recombination with the DNA fragment from (a)
- (d) culturing the cells from step (c),
- (e) selecting the cultured cells from step (d) for the presence of the heterologous sequence, in particular the selectable marker,
- (f) producing chimeric non-human mammals from the cells of step (e) by injecting these cells into mammalian blastocysts (preferably mouse blastocysts), transferring the blastocysts to pseudo-pregnant female mammals (preferably mouse) and analyzing the resulting offspring for a change of the T, T2 or T3 gene.

27. Transgenic cell or tissue which is capable of expressing a T protein or part of the T protein or an ortholog thereof.

28. Use of the non-human mammal according to any of claims 21 to 25 or the transgenic cell or the transgenic tissue according to claim 27 for the analysis of the function of the T gene family.

29. Use of the non-human mammal according to any of claims 21 to 25 or the transgenic cell or the transgenic

tissue according to claim 27 for identifying inhibitors and enhancers of the T gene family.

30. Vertebrate gene and functional equivalent, derivative or a bioprecursor thereof, which code for a protein having a statistically significant amino acid sequence homology to the T gene, T2 gene or T3 gene according to any of the following figures: figure 1, figure 9, figure 11, figure 12, figure 13, figure 14, figure 15, figure 16, figure 17, figure 18 or figure 19.
31. T gene and its vertebrate orthologs and vertebrate paralogs which code for a nuclear pore protein.
32. Vertebrate protein which has an amino acid sequence according to figure 1 or an amino acid sequence which differs from the amino acid sequence in figure 1 by one or more amino acids.
33. Vertebrate T, T2 or T3 gene and the protein encoded therein in all of its naturally occurring allelic and mutated forms.
34. Medicament containing a protein according to claim 9 or a functional equivalent, a fragment or a bioprecursor thereof in combination with a pharmaceutically acceptable carrier.
35. The method of identifying substances which has an enhancing or inhibiting influence on the effect of T protein, T2 protein or T3 protein, by means of
  - determining the bi-directional transport through the nuclear pores,

- determining the binding to filaments of the cell (e.g. actin filaments and microtubuli) or
- determining the increased or reduced transcription of cellular or reporter genes.

36. Method of identifying substances which have an enhancing or inhibiting influence on the effect of proteins which are functionally linked to the T protein in direct or indirect way, or represent parallel signal or functional pathways, by means of

- determining the bi-directional transport through the nuclear pores,
- determining the phosphorylation and the dephosphorylation of proteins,
- determining the binding of the T protein to filaments of the cell (e.g. actin filaments and microtubuli), or
- determining the increased or reduced transcription of cellular or reporter genes.

37. The method according to claim 35 or 36, wherein the modified transcription with reporter molecules, preferably the occurrence of certain mRNAs or the EGFP protein, is detected.

38. The method of identifying further proteins which play a role in the development and function of the nervous system and/or are a nuclear pore protein, wherein the method comprises the steps of:

- (a) producing an antibody against a protein according to claim 9,
- (b) contacting a cell extract with the antibody and identifying the antibody/protein complex,

- (c) analyzing the complex to identify a protein which has bound to the protein of the complex and is no antibody, and
- (d) optionally repeating steps (a) to (c) to identify further proteins of this function.

**Abstract of the Disclosure**

The invention relates to a protein (TP) and to proteins related thereto, which are involved in the development of the nervous system, especially the central nervous system, and are expressed in a tissue-specific and development-specific manner as well as to DNA sequences coding for these proteins. The invention also relates to antibodies directed against these proteins or fragments thereof and to anti-sense RNA or ribozymes which are directed against the expression of said protein. Finally the invention concerns medicaments and diagnostic processes in which the above mentioned compounds are used. The invention further relates to a non-human mammal whose TP-coding gene is modified.

Human cDNA sequence

Fig. 1

60/14549

Fig. 1 (cont'd)

Fig. 1 - (cont'd)

Fig. 1 (cont'd)

## Human genomic sequence

1 GATCAGACTT TGAAGAGTGT TTGTACCAGT CTAAAGTTA CAGAATTAT  
 51 TCCTGCTCTT TGAGGGTGCA TTGCAAATCC AGGCTAGAGG GAGAGATAAC  
 101 AGTTAGGAXA GTACAGCAAT ACTCTACTGG GAAATGGTGA GGTGTTTCGT  
 151 GAAGACAATG GCAACACAGA TGAAGACATG CAGATGGAGG AAATAAAGAT  
 201 CCAGTTGAGC TTGTTGCCA GTTGGATAGA GGTTGAGGTT ATGCATGATG  
 251 GAGCAATCTA GGTTTTGTC TTGGGTAGGT GTTTCCATGA TAGTACTCAG  
 301 AATGAATCAT ATAGTTGTAC AGGTTGAATC CCACCCATGT TTGCACAATA  
 351 GAGTGACTGT CTAGCTGAAA TCCAGATGAC ACTCTGTATG CTAAGCTATG  
 401 CTTCATGGAA CTGTATAAAG GCACCTGCTA CATAGGCTAG TGGCAGATCT  
 451 GGAAGTAACC TATATGGTAT ATAGGAAATG AGGTGGCTTT TGTATAAATC  
 501 CTACAGATAA ATTTCATTTTC CTGATCCTAT TATTTGACT CATGTTAGCC  
 551 CAAGAAGAGT ATTCACTACT TCATATCCCT GAAGGTAAGA CAGAGTAGTA  
 601 TTAGATTAC TATTGGCAA ATAAAAGGGA TCAAGTCCTA AGATCAAGCT  
 651 GATGAATCAA CACCTCATAG GATATGTCCC AACCAATTAT ATGGCTTCCC  
 701 CTATAAATAA AATCTAGTTC TCTTCTCTGG AGAGGAACAG TGAAGAATAT  
 751 CATAACCTAT GCTACAAACT GCTTGAGTAG GAGCTACTTC TCTCCAAGGC  
 801 TTTATATCAT TCATTCTGGC AGGCCCTCT GTTGTGTTCTC ACCAGCTCCT  
 851 GGGAAATTAA TTTCTCCTCT AGTGATATAA AAGCTCTCTG TTTGAGATGA  
 901 AGGGCTGCC AGTTTATCAG ATCTGTATTA GTCTGTTCTC AGGCTGCTAA  
 951 TAAAGACATA CCTGAGACTG AGTAATTAT GAAGGAAAGA GGTTTAATTG  
 1001 ACTCACAGTT CCACATGGCT GGGGAGGCCT CACAATCATG GCGAAAGACT  
 1051 AATAAGGAGC AAAGTCACAT CTTACATGGC TGCAGACAAG AGAGCATGTG  
 1101 CAGGGGAACG GCTCTCCATA AAACCATCAG ATCTGTGAG ACTTGTTCAC  
 1151 TATTACAAGA ACAACAGACA GGAAAACCCG CCCCCCTCAAT TCAATTACCT  
 1201 GCCACTGGGA CCCTCCCACA ACACATGGGG ATTATGAGAG CTACAATTCA  
 1251 AGATGAGATT TGGGTGGGG A TACCGCCAAA CCATATGAAG TTCTTTCTTT  
 1301 GTTACTGGGT ACCATATCCA TTCTGTTGAG GTTCTGAGCC TTTCCAGTTA  
 1351 CTGTAACCTCC TCTATCTCCT GTCTGTGCTA AGACTCAGTG ACCTCTCTCT  
 1401 GCCTTGCTTC TGCTTTGTCC TGACCCCTTC TGTGCATGCA CTCACTCTAG  
 1451 TTTGCCACC TGAGGTGAGA GATGGTCCAG ATTAGCAACA ACAATCTGTG  
 1501 GACTAAAATC CTCTTTAGGG AGGAAGCAAA ATTCAAGATGG ATGTTACTAA  
 1551 ACAAAAGCTCA GAAACAGAGA CCAGGGTGTG GGAAGTAAGG TAGTAGCCTG  
 1601 AGAGCAGCTG GCAGTGTTT AGACCTGGAG GGAGGTTAGG TCATCAGCAA  
 1651 TGAGGAGACT GCCTGGAAAA TCCTAGAAAA TTAAGACATC TGGTCAGGCA  
 1701 AGGTCAATATC ACCAGCACAC TTCCCTTTTC AAGTTGAATC CCTTTCCCT

6/ 124

1751 GTTAAGAGGA TTCAAGTGTC TTTCTTGCAT TTTGTCTTCT CTTCTATATC  
 1801 CATGCTTGCA ATATAAGGAG ACAGCAGTTG GCTGTTGTG CTAGAAAATA  
 1851 TAAATGGCCA TTTTGAAAGC ATGCCAGACA GGATCTGCGG CAAGTTTCA  
 1901 ATGTTACTGC TGCCATCTGT TGTTCTTCAG TGCTGGGATG TGAATCTCTT  
 1951 GGCAAACATC TCTCTAATTC TGAACATATCT TTCACCCCCA TCTAGAGATA  
 2001 TTCACTTACT GAAGTGCCTT TTTAAAGCAA TGTTCCCTCAC CAAGGCGATG  
 2051 TTCTGAATGT TTTAAAATGG AAGAATCTGG AATGTTTTA TTATAATACA  
 2101 TTTTGTATAT CCCAAAGCAA AAATCAATT CTTCATGGTT AATACTTTG  
 2151 TAATTTGTT TTTAATAATA TTTTCCTTT AAATATAAGA AATATTTTAT  
 2201 TGAATTAATA CTTTAATGTA GCTGTTCAA GTAAGATAAA ACAGAACAGA  
 2251 TTACTGTTT CAACCTTGTT CACAGTTAGC TCTGTAACTA AGTTGTTGAG  
 2301 CTTTATCTAA GCTTTTTAT TTTTACATAA CGTTCCCTT TTCACTTAAC  
 2351 CTTGAAATTA TAGTAATTG GGAACCTCTA TTCCTCTGAA AGAGAAAGCT  
 2401 AATGCCAAAG ATATTCAAG GGAGAAAGAA GGTTTTAAA AGGAGAGACA  
 2451 ATTCAAGCTCA GACTTAATAG CTGTGATTGC TATTTATTAA GCAGAACGCC  
 2501 TATAACTAAA TTCTCAGATA TCCAAAAAAC AGCCTGTACA TTCTCAAAAG  
 2551 TGAAGATTAC ACATTTCTA AGTTAAGGTA AAAGTTTGT CTCTGTAGCA  
 2601 TCTTACTGAT TTCTATCTTC TCATTCTGCC TTAATAATGT CACTAAATAA  
 2651 ATGTTTGATG CACTAATACA TGAATAAAAC TATTCACTGGT AATGATTCTT  
 2701 TAGAACACAA GCTAAGTTT GTAATTGTT TTTTAAAAA TTAAAAATT  
 2751 AAATATAAAA ATGTTTTAA AAGGCTTGAA TTTCTGTAA AATGTACACA  
 2801 TTTTAAGTTG TAGGCTGTCT TTAAAAATAA TCTCTCCACA CACTGTAGTA  
 2851 TTTAAAACAT CATGATATTAA CTATAAAACA TCAACAAATA GGGCAGTGG  
 2901 AAACATGGTA ATCACTAAAA ATGCTCACAT GTCATATATT AAGACTTGAT  
 2951 AAGTAAACCA CAATAATAAA TAGAAAAGAA ATAGTTGTCT AAAAAGGGAT  
 3001 TCTCACCTT CAAACCTTAC CATAAAATG GAATATAAAA GAAGGAAGAG  
 3051 GAGGAGAAAT CAAATTATAT CATAAAATT TCTGGGCAA AATATTACAG  
 3101 AAGAAAATAA GAAAGATTAA TGGAGTTGAC TGAAACATT TTGAATCCTA  
 3151 TACATAAAAA TATCGTTAAT TAAAAGAAA AACAAAGAAA CAGATTTGGG  
 3201 AAATATTGTA AACTGGTTT TTTTAGCAT TTAAAAATGT AATACAAATG  
 3251 GATTATTAA ACTCCATTGC AAAAATACAC AAAGGACATT GACAATGTCT  
 3301 GGAAATAAAA TTAGCTAAGT AAGTTATAGA AAAACTCAGT CTCACAATT  
 3351 GACAAATGTA ACTGAAAATC ATTAATATAA TTAGTAACTA TTTTTACATG  
 3401 TCAAAATTT TGAATTACTA AAGGAAACCA CAATGCCTGA AAGTATCCAG  
 3451 GGTTTTTTT TTTTTTATA ATATTGGCAC TGTCATATGG GTGGCAGGAA

7 / 124

3501 TTGAAGTGAT GTTGTTCCTT CAGTTATTAA GTTGCATCTG CAGTGTTC  
 3551 AATGTCCAAA ACCTGTGAGT CAGTAATTCT CTTTTGTAT ATTTATCCTA  
 3601 ATACAATAAT TCTAACATA ATCTCAATAT ATATGTACAA AGTTATTAC  
 3651 TGCAGTGTAA CTTACAATAG TTAGAAAATT GTAAAATGCT TTATGCATCT  
 3701 TAAAATATAA ATTGGTGAAT ATATAATAGT CCATATGATA TAATTATATC  
 3751 ATTATTATAA ATAATGAATT AGAAAATAAT TTAAGAGCAT TAAAATAATT  
 3801 ATAAGGTAAT ATGAAGTGAA TGAATAATGT ACAGACTA TAATCAGCAG  
 3851 AGTGTAACT AGGTAAATTT TTATGTGTGT ATATACTACT TCCTAAAAAT  
 3901 GACTTGACAG AAATCATCAA AATGCTAATG GTGGTTACTT CTGGGTGGGA  
 3951 ATACAGATGA TTTACTTTGT TCCTTTATG TATTTCTGCA CTGCCAGTC  
 4001 TTCCACAGTG AGCATATATT GGTTTTAAA TTTATATAAG ATGGAAAAAG  
 4051 ATACAAATG GTCTTCAATG AACCTGGAG TTAACTTCA TGTGTGTCAT  
 4101 ATGTTATATT CTAACATTAT CACAAATAGA AGACTTTAAA TCAACTTGTA  
 4151 CCTATTTCAA CTATATAACA GCATCTTTAA AATGAGCATT GAATTAAACT  
 4201 ACCAAAACCA ACCATCATGA GGATTATTCA AGTAATGTGT TAAACAAAA  
 4251 GAATTTGTA TAAAATTACT TTATCTCCTT TGTGATTCA GCCCATTAA  
 4301 AAAAATAGA TGTTTCTACT CTCCTTCAGA TATCATTAAA ACATAAAACT  
 4351 GTGCCTGACT GCATAAATCC CTTTAAACT AATATCACTT ATTACGTTA  
 4401 ACTAAGTCTA CCTAGGGCTT CCTTGTATAA AGAACAGAG CTTTCCATT  
 4451 TTTGTTTACC TAGCCCTTTC TGATGCCACG ACAGAATAGC TGTAAATCTT  
 4501 CATTATTTAT ATTCTAGAGA AAATAAAAGC AAATAAAAG GTCAGTGTAT  
 4551 AAAGTTTATT GGTTGTTCTC TTTACTCAAACCCACATGG TATTAATGTT  
 4601 AGTCTCTATG AATATTCAT GGATAAAATC AGAGCATTAA GTGCATACTA  
 4651 AAAACAATAA GAATGGAAAG ACTTTAACCT TATGTTTATA TGAATTCTA  
 4701 GGTATCAAG AAGTTTATAG GCTATAGGCT ATAAAGTCTT AGGCTATGAT  
 4751 ATAGTAACCT AATGTAGACT TCCCTTGATA CATGAAAATA ATGGTACTAA  
 4801 GTACAAACAG AAGATGAGCT TAAAATTATT CTTTGAGTCC TCTTGATGGA  
 4851 TTTTTCCCC CACACTTCC CCAAATTGT TTTATGCCTA TATTGTAGGA  
 4901 GACCATGCAA GAGACCTAGA GTCTTTTTT CTTTCATCAC TTTCCAATCA  
 4951 ACAGCAAATC CTATCATTAA TACCACAAAA TATATCTGAA AACTCCCTTC  
 5001 TTTGATTAACTTGTAACTC CCCATCAAAACTGAAGAGT GTCACAATAC  
 5051 TTCATTAAGT TCCCTACTTG CACTCTACCT TTAATATATT TGTAGCACTA  
 5101 AAATGTTTTT AAAACATATA TCTGCTTATG TCATTTACT GCTCAAAACT  
 5151 ATCTGATTAACTTGTCACT TCTAAGATAC TCTAATTCT TAGCACTCTA  
 5201 TATAAAATCC TTTAAGGGCT TCCCTGCTCA CCTTTTCAGA CTCAGAACTA  
 5251 TGTATTTCTT TTTGCCTGCT GTACTTGTAC CACTGGATTC TTGATTTTG

8/124

5301 TTACTTCCAG GTTTTACAC TTATTTTAC AATAAATGTG AAATACCCTT  
 5351 TTTGACAATA TCTACAAATA TTTCTTATTG GTCTTATTG CTCTTCCTG  
 5401 TAATGTTAG TCTTCATTTC CCTGATAATG GCTATCTAAA GTTATCTCCT  
 5451 CAAAGAAGCA GTTATTATT CACCAAATC TTCTAGTCCT TCTCTGGAGT  
 5501 TTTCTTCTCA CTTCATTCCTT TTGGTTTTTG CCACAATTG TAATAATTG  
 5551 CAATTTGGAG TGTTAGAATG AGGAAATAAA TCACAGGTAA TGACTATAGT  
 5601 TTGTGACTAT GTAAGATTGG ATTCGTTATT GATTATTCC ACAAACACTG  
 5651 AGGCACTGCA TTTAGCCAAA TGCCAATCTT GGGCAGTGAG ACTCTGAAAG  
 5701 AGAATCTGCT TCCCCCACCA TAAACTACAA AGTGAACAA CTCAGAATGT  
 5751 ACATAAATTAA CAGAATGAAA GCACACTAGA AGTAACACAA GATGTGGAAG  
 5801 AGGTAAAGTG TCCTTGAAAA TCATGGAAAG ATTCAAAAG GGAATGACAT  
 5851 TTCAACTGGA TTCTAAACCA GTTATTCAAG CTCCACAAGG TTGCACAGTA  
 5901 AATGAGCAGT GGCAGGATGA CATACTTAG AAAGTAAAAG GAATCTTTT  
 5951 TAAACTGCTA TAAAATCAT TACATATACA TTTTGTAGGT CGAGAGTAAG  
 6001 GTATTTAACAA TAAAATCATT TTAGTATATC AGTGTATATA TAGACTTAGG  
 6051 TTTTCTCAT TTAAAACCTC TTTTAATGAC TTGTGCTTT CTTCATGGTA  
 6101 ATAAAACATT TTCCCAGGAA GTGCTGAATA AATCTTCTT GAAATACGTT  
 6151 TTATTGCTTT CTATCAATGA CCCTGAAGTA ATACAGAATT TACACTTCAG  
 6201 CGGTTGCAAT GCTAAACATT GACAGGTAAT GCACTGTGTT TGCTGATATA  
 6251 AGAGGTATGA TGTAGGGCTA AGTGGTTTG TGCTCATTTA GCTTTCAGGA  
 6301 GAAAATAATT GACTAACAT TTTGATACTA AAACCCAAAG CCTAACAGTT  
 6351 AATTCTTGGT ATTTAAATT ATTATTGCAA AGATTATTGT GCCGAATAAT  
 6401 ATGAAAATAT TTTATATAAT ATTTAAAAAG TATATCTCTT TCTTGGTATT  
 6451 ATTTAAATTAA CCATAAAAAT GTGCGAAAAA GTTATACTGA AATGTGATAG  
 6501 GATCTTTAA AAGTGGTGCC TTGATTTGT TAAGTGTAC CTAGTTTCC  
 6551 TCTGAAAACA AGAAACATAC CCAGAAGTTT TCACGAAATG GTCTCATGAA  
 6601 TATCTAAGGT TAGTCCGTAG TCTCATCTGA GACAAGGAA GTCCCTTCCA  
 6651 CTATGAGCCT GTAAAATCAC AAGCAAGCTA GTTACTTCCT AGATACAATG  
 6701 GGAGTACTGG TATTGGGTAA ACACAGCTGT TTCAAATGGG AGAAATTGGC  
 6751 CAAAATTAAAT GGGTACAGG GCATGCAATT CCGAAATCCA TCTGGGCAGT  
 6801 CAAATTGTAA AACTCCAAA TGATXTCTTT TGACTCCATG TXTCACATCC  
 6851 AGGACATGCT GAXGCAAGAG ATAGGTTCCC ATAATCTTG GCAGCTCTGC  
 6901 CCCGTGGCT TTGCAGGGTA TATCACCCCT CCCAGCTGCT TTCACAGGCT  
 6951 GGCATTGAGT GTCTGTGGCT TTCCCAGGAA CAAGGTGCAA GCTGTTGGTG  
 7001 GATCTACCAT TCTGGGGTTT GGAGGATGAT GGCCCTCTTC TCATAGCTCC

9/124

7051 ACTAGGCCGT GCTCCAGTAG AGACTCTGTG GGGGCTCTGA CCCCAGATT  
 7101 CCCTCCTGCA CTGCCCTAGC AGAGATTCTT CATGAGGGCC GTGCCCTGC  
 7151 AGAAAACCTCT TTCCCTGGCA TCCAGGCATT TCCATACATC TGAAATCTAG  
 7201 GTGGAGGTTTC CCAAACCTCG ATTCTTAATT TCTGTGCACC TGCAAGGCTCT  
 7251 CTACCACGTG GAAGCTGCCA AGGTTGGGG CTTGCACCCCT CTGAAACAC  
 7301 AGGCTGAGCT ATACCTTGGC CCCTTTAGC AATGGCTGGA GTGACTGGGA  
 7351 CACAGGGCAC CAAGTCTCTA GGCTGCACAC AGTATGGCA CCCTGGGCC  
 7401 AGCCCTCAAA ATCATTTTTT CCTCCTAGGC TTCTGGATCA GTGAAGGGTG  
 7451 GGGCTGCCAT GAAGACCTAT GACATGCCCT GGAGACATTT TCCCCATTGT  
 7501 CTTGGGGATT AACACTGGCT CCTTGTTACT TATGCAGATT TCTGCAGCCA  
 7551 GCTGAATTTC TCCTCAAAAA ATGGGTTTTT CTTTTCTACT GCATTGTCAG  
 7601 GCTGCAAATT TTCTGAACTT TTATGCTGTT TCCCTTTAA AATGCGATGC  
 7651 TCTAACACA CCCGTCACCT CTTGAATGCT TTGCTGCTTA GAAATTCTT  
 7701 CTGTCAGATA CCCTAAATCA TCTCTCTCAA GTTCAGAGTT CCACAAATCT  
 7751 CTAGGGCAGG GGCAAAATGC CACCAAGTCTC TTTGCTAAAA CATAACAAGA  
 7801 GTCGCCTTG CTCCAGTTCT CAGCAAGTTC CTCATCTCCA TCCGAGACAA  
 7851 CCTCAGCCTG GTCCTTATTG TTTATATCAC TATAAAAATT TTTGTCAAAG  
 7901 CCATTCAACA AGTCTCTACT CCAAACTTTC CCACATTTTC CTGTCTTCTT  
 7951 CTGAGCCCTC CAAATTGTTTC CAGCCTCTGC CTGATACACA GTCCCAAAGT  
 8001 TACTTCCACA TTTTGGATA TCTTTTCAGC AATGCCCGC TCTACTGGTA  
 8051 CCAAACTTACT TTGTTAGTCC GTTTCACAC TGTTGATAAA GACATACCCA  
 8101 AGACTGGAAA GAAAAAAAGG TTTAATTGGA CTTACAGTTC CACATGGCTA  
 8151 GGGAGGTTTC ACAATCATGG CAGGAGGCAA AAGGCATTTC TTACATGATG  
 8201 GCAGCAAGAG AAAATGAGGA AGATGCAAAC GCAGAAATCC CTGATAAAAC  
 8251 CATCGGACCT TGTAAGACTT ATTCACTACC ACTAGGACAG TATGGGTGAT  
 8301 ACCACCCCCA TGATTCAAAT GATCTCCAAAC CAGGTGCCTC CCACAAACACA  
 8351 TGGGAATTAT GGGAAACAA TTCAAGATGA GATTTGGGTA GGGACACAGA  
 8401 GCCAAACTAT ATCACATGGA TTTCTTATAC TTTGCTTTT AATAACACAA  
 8451 ACAAAAAAAAT ACATCATTAA AAGGTTAGAA GTGAGAAGGT GTTTTTATGG  
 8501 AAATCAAAAA TAATATCACC TTAGTGAACA GTATTCTTAT GATTGTAGTT  
 8551 GAATTAGAGA GCAGAAATACA TCTAGAAGAT TCAGTAGTAA GCATGTTCT  
 8601 TCGATTAATG GAAAATTGTA ATAGCCTAGC TGATTGAGAT TGAGGTTACT  
 8651 ATTAAATGCC TGAAGTATAA GAGTTGGTTG TTTATGTAAA CAAAATATCT  
 8701 GTTTTACATG TACATGTGTA AGTAGGACTG TTGAGCCCCA GTAACATGAA  
 8751 ATATCAAAGA GCATGACTCG AATACCTGCC ATATGAAGTG CTATTACATC  
 8801 AAAAAAGAGG CGTGTGCTGA AAAATTACCT ACAAAATGGCA TTTTCCTCAA

10 / 124

8851 ATCAATTTA AATCTTCAGA ATTCATTAAATAATTGTT TAGTTAATAT  
 8901 TTCAGAATCC CTCATCATAA AAAGCAGGCA AAAGGCAAAA GTCCTTGAAT  
 8951 GTATAACACA TTTGTTTCA ACAAGCCTG CCTCTAACTG TGAAATCCAGG  
 9001 AGTGAATCCA GAACTACAAA TTAACTAAGA TTGGCCCCAT CGAGTTACTG  
 9051 AACGTTAAAATCTAAAAC TAAAAGGCAT GCCTCAACAA TTATTTCTT  
 9101 CTTGGAATCA TTAATTAACC TATGTGTATC CAAACAATAA TCTTCCAGCA  
 9151 GTTTCGCTAG CTACATTTT AATTACTTAA TATCATGTAA AATTGTTTT  
 9201 ATTATTGTTCA AGTTCTGAAT TTTGACATAT GCATCAAGCC ATGCAAATGCC  
 9251 TACCACAGTC TCCCTGATCA CTGATCTGTT CAAATCTCT ATAGCATTTC  
 9301 TCCTTTCTT AAATGTTGCA TAAATAAAC CATAACCTTAT GTGGCCTTTT  
 9351 GAATCTGGCA TCTTAACTT AATGCGCTTG AAATTAATCT ATGTCATTTC  
 9401 ATGTATCAAT GGCTCAATCT TTTAATTGT TAAGAAAAAA TGTATGCTGG  
 9451 GATAAAATATC TTTCTAAATG AGTTTTGTT CACAATGCTG AGTGTGTT  
 9501 TAGGATAGAG TCCTAGAAAT GGTATCACTA GGTCAAACAT TCAAATAATT  
 9551 TAAATATT TGATACATAT TGCCAAATAA TCTCAAATT TTTACCAATA  
 9601 TACATTTATG ACAGTATGGG ATAAATGTGT CTTTCTTATA CCAACTGACA  
 9651 ACATTAATGA TAATACATAA AATATTCTT GCTAATTGTA TGGGACAGAA  
 9701 ATGTTATATC CTTATTAGCA TTTTATTATT GTGGTGAAT GACTGTACTG  
 9751 TACAGCCAGA GATATTGGT TCAAAATCCA TCTTCATTAT TTACTGTATG  
 9801 TGAAAATTAA GGTGAGCTAT TTAATCTCTT GATGCCCTAG TCTCCTAATC  
 9851 TATAAAGTGG GGATAATTGT ACCAATCATA TTAGGTTCCCT GTGAGAATTAA  
 9901 ACTGAATTAC TATAGAAAAT GCTTAGAATG GTATCTAGTC ACCAGGAAGG  
 9951 ACTCTCTCTG TATTACTTGT TTATTATCTA ACACGTTAA TTATTAATGA  
 10001 AGCTCAGTTT CGTTATATGC TTGGGATATT TGAAACCTTT CTTAGTGAAT  
 10051 TTTCCAATAA AATTATTGT CTATTTCT ATGGACAAGT TGGTATTATT  
 10101 CTTACTGGTT TGTTTCAGGT TCAGTTAGTA AGAATTAA GGATTTCTA  
 10151 TCACATTTA GCAAACCTTT TCTGCATTTT ATCTTTTTC TTTCAGATAA  
 10201 TGTTTGCAAA ATGTAAAAAA AACAAAAGGT TTCTTCATCA AGTTGGTATC  
 10251 TTTATCTTTT TTATTGCTTT GTGATTGAA AATTCTGTC CTGAGAACCA  
 10301 AAATATATAT TTGATGAAAT AGTTCTCTC TTTTACTCAT TCTGAAGTCA  
 10351 TTGGAATTGA ATTTGGCATA TGATATAAAT CCTAATTAA TATTTTATGA  
 10401 TATTCAAAAT TTCTAACAAA TATTACTTA ATAATCTAAT CCAGGTTCT  
 10451 ATTGTTCTT CTGTTCCCTT TATAATGCTT TTTCTGAAGT TATTTTCCT  
 10501 AGACTTAAAT ATTAGTATAA TATTATCATA GAGGAAAAAA TATCTGTTAG  
 10551 CTATGAATAA AAGGCTTTCA TCTTATTGTT GCATTAATAT ATTTAAATGT

10601 AGAGAGCATA CAGATTAGCA AAGAAAAAGT ATAATGCCT TTTTTTATAG  
 10651 TTGACATGAA CATGTATAAA GAAAAACCAA AAAATCAAT AAAACAACCA  
 10701 GAACTTATTA GTGAATTAG CAAGATCATA GCATACAAAG CCAAGATTCA  
 10751 AAATTCCATT TTATTTATCT ACTAACAAAA AATATTGAA ATTTGAAAAT  
 10801 TTAAATATGC CATTACAAT AACATCAAA TATTGAACAA TAAAGTATT  
 10851 AGGAATTAT AAAATGAAAT CTCCTATACC AGGAATTACA GACCATTGCT  
 10901 GAAATAATG AAAGAAGACC AATATATGTG AAGAGATACT CATTGTGGA  
 10951 TTGAGAGACA ATATTGTTAA AGTATCAGTA TTTCCAAAT TAATCAATAG  
 11001 ATTCAATATA ATGGTGAACA GAACACCAGA AGATGTTCTG TCGAAGCTGA  
 11051 CAAGCTATTT CTATAATTCA AATGGAAATG CAAAAGGCAG TCACTGCCAA  
 11101 CACCAGCATG GACTGTCCTGG GTTCCAGTAG GTTACTTCAC TACTGCCTCT  
 11151 TCTGTCAGCC ACATCACGAC AGCTGCCAG AAGCCAGAGA AACTCCTCAC  
 11201 ACCTGGCCA CTGCTGCAGC TACCAAGCAGC CAGGCAAGCC ACCATCAGCC  
 11251 CACTGGTAAC TGCCAACAGA GGTACCACTG TACACTACCC TGGGAAACAA  
 11301 AGATAGGCAT GTAGTCAGCC CACCTCTGCC ACCACTAGGG CCTGAAGCCT  
 11351 GGCCCACCTG ACACTGCAGT CCTCAGCACA GCTTCATCAC AGCTTCTGTT  
 11401 AATAACCACA CCCTAACCTA CCAAGGAAAT CACAAATGTC ACTGACACTG  
 11451 TTTGTAGCCA AAGAAATCAT AGAGAGACTA CATTACTGCA CACACCCATA  
 11501 ATCAAAGCCA CAGTACCTA TCCAGACAAAC ATCACAGGT AATCTAAAGG  
 11551 AAAAAATTTC CCCATATGAA AGCGAATTCA AATATAGGAA GAAGCGACTG  
 11601 TTACAACAGA TATGCAGATA AAGCTTCAAC AATATCCTAC ATTCAACCAG  
 11651 AAGAAAAGAAT CTCAGAAGGT AAAGACAGGT CTTCTGAAAT AATCTAGTCA  
 11701 GACAAAATTAA AAAGAGAATA ATCAAATCCT TCCTGACATT TGGGATAACA  
 11751 TTAAAGTGAC CAAATATACG AATTATAGAT ACCCCTGAGA GTGAAAAGAC  
 11801 AAAGAAAAGA TTAGAAAACC CACTTAATTAA AATAATATAT GAAAACCTCC  
 11851 TAAGTCTAGC AAGAGTTTA GATATTGGG ATGCAGGAGG CTCATGGTC  
 11901 CCCAGGCCGA TAAAACGCAA AAAGGTCTTA TACACAGCAC ATTACAATCA  
 11951 GACTGTTAA AGTCAAAGAT AAGGAATAAA TTCTAAAAC AGCAAGAGAA  
 12001 AGTGTATGAT AACCTATGAA GTAAACCTTA TCAGACTGAC AGCAAATTTC  
 12051 TGGCAGAAC TTTACAGGCC AGAAAAGAATA GGACAATATA TTCAGGAGC  
 12101 TTAAAGAAAA AAAAAACTAT CAGCCTAA TACTATAGCC CACAAAATTAA  
 12151 TCCTTCATAA ATGAAGGAGA AATAAAAGGT TTCCAGACCA CGAAAATGCT  
 12201 GAGGTAGTTT GTTACTACTA GACTGGACCT ACAATAATG CTCAGGGAG  
 12251 GTCTGGAAAC TGGTAGTGAA AGGACGACAT TTATCATCAT GAAAATACAT  
 12301 GAAAGTATAA AACTCCCTGG TAAGCAACTA AAGGGAGGTA TCAAATGTTA  
 12351 CCACCAGAGA AATCTAACTA ACCACAATGA CAAACAATAA GGGAAAAGA

12401 AAGGAACAAA AATATATAAG ACAACAAATA AACACAATA TAACAGGAAG  
 12451 CCTCACATAT CAGTAATCAC TTTGAATGTA AATGAATTAC ATTCTCCACC  
 12501 TAAACGTTAT GAAATGCCTG AATGATAAAA CTATATGATC CAAATATATG  
 12551 CTGATTACAA GAAACTTACC AGGCAGACAT ACATAGGCTG AAAGTAAAAG  
 12601 AATGGTAAAA GATATTCCCTT GCAAATGGAA AGCAATAGTG AGCAGGAGTA  
 12651 GCTATACTTA AATTAGATCA TACAGACTTT AAGTCAAAAA GAGTAAAATA  
 12701 AAAAAGACAA AGGATGTTAT TATATAATGA TGAGATTAAC CCAGCAATGG  
 12751 GAAATAACAA CTCTAAATGT ATATGCATTC AACACTAGAG AACTCAGATC  
 12801 CACAAAGCAA ATATTAGACC TAAAGAGAGA AATAGACTGC AATACTGAA  
 12851 TAGTGGAGAA CTTCAACACT CCACCTTCAG TATTAGACAG ATAATCTAGG  
 12901 CAAAAAATCA ACCAGTAAAT TTTAGATTTA AACTAGATTT TAGACCAAAT  
 12951 GGACCTAACAA GACATTTACA AAACATTCCA TCCAACCACT GCAAAATGAA  
 13001 ATTTGTGTCA TCAGCACATG AAACAATGTC CAAGATAGAC CACCATATGT  
 13051 TAGGCCACAA ATCATGTCTC AGCAATTTTT TAAAAGTTGA AATCATATCA  
 13101 CATATCTCT CAGACCACTG TTGAATAATG CTAGAAATCA ATGCCAAGAA  
 13151 TAACGTTGGA AACTATACAA ATACATGCAG ATTAAACAAAC ATGTTCTGG  
 13201 TTGATCACTG GGACAATAAG GAAATTAAGC TGAAAATCAA AAAATTCTTG  
 13251 TAACAAATAA AGATTGAAAC ATAACATATC AAAACCACTG GCATACAGCA  
 13301 AAAGCAGTGC TAAGAGGGAA GTTTATAGCA ATAAATGCTT ACACTGAAAA  
 13351 AGTAGAAATA TTTTAAATT AGCAACCTAA CAATGTGCCT GAAGAAACTA  
 13401 AAAAATCAAG AACAAATCAA ACCCAAAATC AGCAGAAGAA ACACAAAAAT  
 13451 AAAGATCAGA AAAGAACTAA ATCAAATAGA GACTAAAAAA ATACAAATGA  
 13501 TTAACAAAAC TAAAATTGG TTATTCAACA AGATAAATAA AATTGATAAA  
 13551 CCGCTAGATA GACTAAACAA GGAAAAAGAA TATCCAATAA AACACAATCA  
 13601 AAAACGATAA AGGAGACATT ACAACAGATG CCACAGAAAT AAAAAGGATC  
 13651 ATCAGAGACT ATTATTAACA ACTATATGCT GAAAAATGGA AAATATAGAG  
 13701 AAATAGATAA ATTCTCTAGAA ACTTACAACC TACCAAGCTG TTGCATCAGG  
 13751 AAGAAATAGA AAACCTGAAC ATATCAGTAA TGATTAGCAA AATTGAATCA  
 13801 GTAATAAAAA ACATCTCCCAC ACTCTTTAA AGCTTGGAC CAAATAGCAT  
 13851 CACAGCCTAA TTCTACCAAT CATGCAAAGA AGAATACCAAG TCTCTTGAT  
 13901 GCTATTACAA TAAATCAGAG GAAGGAATTC TCTCTGGCTC ATTCTACATG  
 13951 ACCAGTGTCA CCTTGAAACC AAAACCTGAC AAGGACACCA CAAAAAGAAA  
 14001 ACTACAGGCC AATAACCATG ATGAACACAG ATGCAAAAT CATTAAACAA  
 14051 ATACTGGCAA ACGGAATCCA ACAGCACATC AAAAAAATAA TATACCACAA  
 14101 TCCAGAGGGT TTGTATCAAG GATACAAGTA TGACTCAATG TAAATAAATC

14151 AATAAACATG ATAAGCATCT TCACAGAATA TAAGACAAAT GAATATATGA  
 14201 TCATCTCAAT AGATGCAGAA AAAAATTGGT GATAAATTTC AACATCTCTT  
 14251 CATGAAAAAA ATCTCTAAA CTCAGCATAG AAGAACATA CCTCAATATA  
 14301 ATAAAGGCCA TATGTGACAA ACTCAGAGCT AATATCATAAGAATGGGC  
 14351 AAAGTTAAA GACTTTCCTC TAAGAACTGG AACAAAGACAA GGATGCAAAC  
 14401 TCTCACCACT CCTATCCACA TAGTACTAGA AGTCCTAGCC AAAACAATCA  
 14451 GACAAGACAA AGAAATAAAA AGTATCTAA TTGAGAAGAG CAAGTAACAT  
 14501 TGTTCCCTTT TGCTGATGAT ATGGTTTGAT ATCTGGAAAA TACTAAAAAC  
 14551 TCCAGCAAAA ACCTCTTAGA TTTGATTAAT TAATTTAGTA AAGTTTCAGG  
 14601 ATACAAAATA AAAATACAAA AGTCAGTAGC ATTTCTATGC CCCAATAATA  
 14651 AAATAGCTAG GAAAGAAATC AAGAAAGTGA TCCCATTAA ATTAGCTACA  
 14701 AAAAATTAAA ATACCTGGGA ATAAATCAAG GAAGTTAAAG ATCTCTGCAC  
 14751 AAAACTACAA AACACTGATG AAAGAAATTAGGATTAAC AACAAATTG  
 14801 AGAAACATCC CATGTTTATG GATCAAAAGA ATTAATATCA TTAAAATGAC  
 14851 CATACTTCCC AAAGCAATT CCACATTCAA TGCAATTCT ACCAAATTAC  
 14901 CAATGTCATA TTTCATAGAA TTAGAATAAT CCTAAAATTAGTGAATG  
 14951 AGAACAGAGC CCAAATAGCC AAAGCAATTG TGAACATAAA GAACAAATCT  
 15001 GGTCCCTGACT TAATCACTAT GCAATCTATG CATGTAACAA ATTGAACAT  
 15051 GGATTTTATC AATTTGTACA AATAAAAAAA TGTAACAGCTTAAAGCT  
 15101 GGAGGCTATA GTAGCCAAAA CAGCATGGTA TTTTAGACA AATGGAATGG  
 15151 AATAGAAAGC TCAGAAATAA AGCCATATAAT ATATATTGTG TGTGTGTG  
 15201 TGTGTATACA CACATACATG TATATATAAT GTGTACATAT AATGTTTCT  
 15251 ACATGTTCTA ATATTTATAT TCCATTCCAT TATACATATT CCATTCTGT  
 15301 ATATAGGTTA TATAGAATTG GAAGACTATC TGCCATTAAA AAGAATGAAA  
 15351 TCCTGTGATT TGCAGCAACA TGGTTGAAAC TGGAGTTCAT TATCTTAAGT  
 15401 GAAATAATCT AGGCACAAAA AGATAAAATAT CACATGTTCT CACTTATATG  
 15451 TGGGAGCTAA TAACTTGATT ACATGAAGGT GGAGAATGGA AAGGTAGGTA  
 15501 GGAAACAGAG ACTGGAAAGG ATGAATGGAG GGTAGGAGGG AAGGTGAAGA  
 15551 GAAGAGAGTT AAAAGGTGTA AACATATAGT TAAAAGAAAT AAATTCAATG  
 15601 CTTGATAGCA GAGTACAGTG ACTACAGTTA ACAAAATGTA TTATACTCAG  
 15651 GTGATGAACA CCTAAATACT TGATCACTAT GCAATTATAT ACGTGTAAACA  
 15701 AAATCACTAT GCACTATATA CGTGTAAAAT TAAATGCGTA CAAATAAAA  
 15751 TAATAAAAATA CTAATCCAGT ATCATTCACT GACAATGTTA ACTCAGGTGG  
 15801 ATAGGCATTA AGTCAATACT ACTATAAGAA CCACTTCTTG TTTATGTTAA  
 15851 TGCCATATAG AATGAAATAA AATTCACTAA AATCCAAAAA ATTAGAAAAA  
 15901 CTATCAAAAC TCAATAATAT TAAGACAAACC CAATAAAAAT GTGGTCAAAG

15951 GATTTGAACA TACATGTCAC CAAAAAATAT ATTCAAATTT CCAATAAATA  
 16001 CATGTAACAA TGTCGACAT CGTTAGTCAT CAGAGAAATA CAAAATAAAA  
 16051 TGGTAATGAG ATACTACTAG ATAGGCCTTT ACAGAGACTG ACAATACCAA  
 16101 GTATTGACAA GGATATGGAG CAACTGAAAT TCTCATTCCCT TGTGGTAAGA  
 16151 ATGTACAATT ATATAACCAC ATTGAAAAAA CAAGTTTCA GTTTCTTTAT  
 16201 TCACCCAAA TATATGTCTT TTGGAAAAAA TTTTTCCAG TCTGTGGGTT  
 16251 GTCTTCTCAT TCTCTTGATA TATGTCTTT CAAAGAGGCT GAGCTTTACT  
 16301 TTAGACAGTG GTCATCAAAG TGTGTATATT TGTGTTTTA TAATTTATAT  
 16351 GCATATATTC CTGTGAAAAG ATACTGTATG CATTGTTCAA CATGTACAAA  
 16401 TATAAGAAAG ATATAGTAAA GAAATATATA TTTCTAAATT TATAAATGTA  
 16451 TTTATTGGTG TTCCACGTTG CAAACTAAAT AATCTACGTT GGCTAATTAA  
 16501 AGGAATTAAA CTATAGTAGA AGGTTCTCAT TTATTGGGAT GATTAGAACCC  
 16551 AGCCTTTTG CAGGCTATTA GCGAATCATA GCACTAGGGC TTCACTGCTA  
 16601 CCTCCACTGA CACCTCTGAC ACTTGAAACT TGAGGCCAGA TATCTGCCCA  
 16651 TGCTGATAGA AAACAACGTGA ATAATTTAAT TTGCTAGATA ATAGAAAAGA  
 16701 ATCAAATGAC TCTGCCACAT TGCTTGCCAG AAGATTGTTT TTCTCATTG  
 16751 TGACCTCTTG CCTATAAAATG ATAGATAGTC CCTGTGCTGC ATGCTATAGG  
 16801 TGTTCGTAAG AGAGTCTGGG AATGTGAGCT TTTTATATCC TATTTTG  
 16851 TGGTAAAGGT CATTCTATTA GTCTGTTCTT AAACTGCTAA TGAAGACATA  
 16901 CCCCAAATG GGTACTTTAT GAAAGAAAGA GGTTAATTG ACTCACAGTT  
 16951 CAACATGACT GGGGAGGCCT AAGGAAAGTT ATAATCATGG GGGAAAGGGGA  
 17001 AGCACACATG TCCTTCACAT GGTAGCAGGA AGGATAATGAA GTAAAAGGGG  
 17051 GAAAAGCCCC TTATAAAACT ATCAAATCCC ATGAGAACTC ACTCTCACAA  
 17101 GAACACAATT AGAGTAACGTG CCCCCATGAC TCAATTACTT CCCACCAGGT  
 17151 CCCTCCCACA ACACATGGGG CTTATGGAA CTACAATTCA AGATGAGATT  
 17201 TGGGTGGGGCA CACAGCCACA CCATTTCATT CCACCTCTGA CCCCTCCCAA  
 17251 ATCTCGTGTTC CTCACAATTC AAATACAATC ATGCCCTTCC AACAGTCCCC  
 17301 CCAAAGTCTT AACACATTTC AGTATTAACA CAAAAGTCCA AGTCCAAAGT  
 17351 CTAATCTGAG ACAAGGCAAG TCCCTCTGC CTATGAGCCT GTAAATTGCA  
 17401 AAGCAAGTTA GCTACTTCCT AGATACAATA GGGTCACAGT CATTGGGTA  
 17451 ATACACACAT TCCAAACGGG AGGAATTGAC CAAAACCAAG GGGCTACAGG  
 17501 CCTCATGGAG GTCCAAAATC CAATAGGGCC ATTGTTAAC CTTAAAGTTT  
 17551 CAAAATTATC TCCTTTGACT TCATATCTCA CGTCTAGGTC ATGATTATGC  
 17601 AAGAGGTGGG CTCCACAGC TTTGGGCAGC TCTGCCTCTG TGGCTTGCA  
 17651 GGGTACAGCC CCACCTCCAGG CTGCTTTAC AAGCTAGTGT TGAGTGCCTG

09/914549

15 / 124

17701 CAGCTTTTCC AGGCACATGG GTGCAAGCTG TAGGTGGATC TACCATTCTG  
17751 TGGTCTGGAG GATGGTGGCC TTCATCTCAC AGATCCACTA GGCAGTACCC  
17801 CAGTGGGAC TCTGTGTGGG GGCTCTGATC CCACATTCC CTTCCACACT  
17851 GCCCTAGCAG AGGTTCACCA TGAGGGCTCC ACCCCTGCAG CAAACTTCTG  
17901 CCTGAACATC CAACCATTTC CTTACATCCT CTGGAATCTA GGCGGAGGTT  
17951 TCCAGACCTC AATTGTTGAC TTCTCTGCAA ATGTAGGCTC AACACCCCCAT  
18001 GGAAGCTGGC AAAGCTTGGG GCTTTCACCT TCTGAAGCCA TGGCCTTAGC  
18051 TGTACCTTGG CCCTTATTAG TTAAAGCTGG AGCAGCTGGG TTGCAGGGCA  
18101 CCAAGTCCCT ATGGTGCATA CAGCAGGGG GCCCTGGACC CAGCCCACAA  
18151 AACCAATTTC CCCTCCTAGG CTTCTGGGCC TGCGATGAGT AGGGTTGCCA  
18201 CAAAATGTC TGACATGCCT TGGAGACATT TTCCCTATTG TCTTATTAAG  
18251 ATTTGGCTCA TAGTTACTTA TGCAAATTTC TGCGAGCAGGC TTGAATTCT  
18301 CCTCAGAAA TGAGTTTTC TTTTCTATGG CATCATCAGG TTGCAAATT  
18351 TTAAAACTTT TATGCTCTGC TTCCCTTTA CAATTAAGTT CCAATTCCAA  
18401 ACCATATCTT TCTGGATACA TAAAACGTAA TGCTTATAAAC AGCACCCAAA  
18451 TCATATCCTG AACACTTGC TTCTCAGAAA TATCTTCTAC CAGATACCC  
18501 AAATTATCGC TCTCAAGTTC AAAGTACCAAG AGATCTCTAG GGCAGGGGCA  
18551 AAATGCCACC AGTCTCTTG CAAAGCATA ACAAGAGTCA CCTTGCTCC  
18601 AGTTCCCAAC AAGTCCCTCA TCTCCATCTG AGACCACCTT AGCCTGGATT  
18651 TCATTGTCCA TATCATTATC AGCATGTTGG TCAAAGCCAT TCAACAAGTC  
18701 TCTAGGAAGT TTCAAACCTT CCCACATCTT CCTATCTTT TCTGAGGCCT  
18751 CCAAACGTGTT CCAACTTCTG CCTGTTACCC AGTGCAAAG TTACTGCCAC  
18801 ATTTCTGGGT ATCTTACAG CAGTGCCCCA CTCCTGGTAC CAATTTACCA  
18851 TATCCATTAA TTCTCATGCT GATAATAAG ACATACCCAA GGCTGGTAG  
18901 TTTATAAAGA AAAAAGAGGT TTAATTGACT CACAGTTCAAG CATGGTTGGC  
18951 AAGGCCTCAG GAAACAGAAT CATGGTGGAA GGGAAAGCAAA CACATCCTCC  
19001 TTCACATGGT GGCAGGGAGA AGAATGAGCA AAACGGGGGA AAAACCTTA  
19051 TAAAATCATC AGATCTCATG AGAACTCACT CTCTTGAGAA CAGCATGAGG  
19101 GTAACCATGT CCATGATTCC ATTACCTCCC AACGGGTTCC TCCCATGACA  
19151 CGTGAAGATT ATGGGAACCA CTACAATTCA AGAGGAGATT TGGGTGGGG  
19201 CACAGCCAAA CCATGTCAGT CATGATATGA GAAATTATCA AATTAAGATG  
19251 TAGGGAAGGT TTTTAAAGA TTTGAGCAAC CACAAATGAC AGATATGTGC  
19301 TATAGTAGTG CAAAATACCA TTTTGCTCTT ATTAAAAATA TAATTGTTCT  
19351 TGATAATCTG AATTATAAAT GTCATGGATA ATTATGATGC ATTATGCTCT  
19401 CAGCAGCTAA AACTCAAGC AAAATACACA CCTAGAGAGC AATCAGCCTT  
19451 AACAAATAATT CTATAAATT AATTCTTCTTT ATTCTGATA ATTACATTAA

Fig. 2 (cont'd 10)

09/914549

16/124

19501 AGTTGACTTC ATATGTGATC TAAATACATT ACCATTATTT TGGACTTATG  
19551 ATGTAGCTCT TGAAGTACAT ATATGATGTA GCTCTTAAAG TACATATAGA  
19601 AGAGCAGATA AAGTATCAGT TCACCATTTC TTTGTAGTTT GTGCTTCAT  
19651 GATGAATATT CTCATCAATG TACAGATTAT TTGCAGGAGC CTTTTAAATC  
19701 CATGTGTCCA TTTTATGAGA CTTAGCTTT GTCTGTATAT AATGTGTTA  
19751 TTCAGTGTGC ATGGATTAAT TTGAGAGAGC ACAGGTATGG GTATCTTAC  
19801 AGCAGTGCCC CACTCCTGGC ACCAATTAC TGTATTAGTT TATTCTCATG  
19851 CTACTAATAA AGACTATATA TCACAATAA CTGAGAACCA GCTGGTAAAT  
19901 GAGAGAACTG TGGTCCACCT TTTCAATTGTG GAGTTCTCAT TTTCCCTAGC  
19951 TTATGCTGCT TATTCAACAC TATTCTGCA TAATCTAATG CATTCACTAA  
20001 ATGAAGGTGC TGTGTTAGCC TCCACATGAT ATTAATACAG CCTATTAAAT  
20051 TTATCCTCT TTAGATTAAA AATAAATAAG TAGTCATGTG CCACAGAATG  
20101 ACACCTCAGT CATTGGTCA TTGAAGGACC ACATCTATTA CTGTGGTCCA  
20151 ATAAGATTAT AATAACATAT TTTTCCTGTA CATTTCATT GTTCTGATAT  
20201 GTTTTGATAC ATAAATGCTT ACCATCGTGT TAGAGTTGCC TGCAGTATTG  
20251 AGTACAGTAA CATGCTGTAC ACCTAGGAGC AACAGGCTAT ACCACATACC  
20301 TTAGGTGTAT AGTTAGGTTA TACCATCTAG GTTTGTATAA GTACACTCTA  
20351 TGATGTTCTC ACAATGAACA AAATCACCTA ATGATGCATT TCTCAAAACA  
20401 TGTCCCTGTC ATTAATACAG TATGTAACAA TACAGTTAGT ACAATATGTA  
20451 ATACATGACT ATATTCAGAA TTTTAGCTAT TTCTCTTATA TTTCAAATGG  
20501 ATTTTCTTAT GCACTGTGTG GCACGGGCAT TTCACTTTAG TAACCACAGT  
20551 CTGGAAAGG AGAAAGTCTT GAAGGATGTT GAGCAAGGTT ATGACATGGC  
20601 CAGATGTGAA TTTTGATCA GTGACTCCAT GTTACCGAGAT AAAGTTGTAT  
20651 TGGGAAAGAT CAAAAGCATG AAGGCCAGAT AAGAGGATAC TGTATGTTAT  
20701 CATGGATGGA AATGTGAGGG ATGGCAGGAG AGATGCTATG ATTGAATGAA  
20751 TCTCAATATT CTTGGTGATC AAAGAATAAT GAGACTCATC CAATAAGACT  
20801 CTGTGAATGA TTGAATGTAG TTCTAAGCT AGGAGGAAGA ATGAGGAATG  
20851 ATTTTCTGGT TCCTGACTAC AGCACAAAGTT TTTGATTTT AGAACAAAGA  
20901 ATAAATTTGT ACATGCTTTA TGATTCCCTGG TTGAATTTT AAGGATAAAA  
20951 AAGTCAGCTG TAATATTATT CTTTCCTGAT ACCATGCAGT ATTTGTATCA  
21001 GTGATCTTAT TCATTCCACA CACATTCTTC TTGAACCTGG ACACGTCT  
21051 AGACACTGAT TCTTCCAAA TATCAGATAA GGTTATTCTT ACGTAGACCC  
21101 TCAGTTCATA TAAATATGAT TTTCCCAAAA TGTGAAATAA GTGACTTTTC  
21151 ATAAGATATT TTTAAAAGA ATGTCTTAAT AATAAATTGT GAATGTTGCA  
21201 TGGAAATGTA GGTGACTTGC ATTGTGCATC CTGTGTTGA TTCACTGCTC

Fig. 2 (cont'd 11)

09/914549

17 / 124

21251 TTGCATGTCT TGCCTTAGC TGGGATGACA GCAGTTCAGT GAGCAGTGTT  
21301 CTCAGTGACA CCCTTGATAA CATCAGCACT GATGACCTGA ACACCACATC  
21351 CTCTGTCAGC TCTTACTCCA ACATCACCGT CCCCTCTAGG AAGAATACTC  
21401 AGGTGAGAAT TACCACCTT CTTTTCCAG TGTTTCTGCC AGCTTTTCC  
21451 CCAAAATTAC TTAATATTAG ATTAAGGTAT AGCACAAAGCC CTTAATCAA  
21501 AATTATTACA GAAACTGGAA AATGCAGAGA TAATAAGGAC TCCCTTTGCC  
21551 ACTCCTGAAC CCTGAAGCAT CTTTCATCTT AGTCTTCCT AAAGCCACAA  
21601 CCCTTAGGAG GAGCAACAAT GTGCACTGCA GCCAATTTG ATAACACAGA  
21651 AGCAGCTTAT ATATATATAT ATATATATAT ATATATATAT ATATATGATA  
21701 TACATTACAT ATTTATATAT ATGTAATATA TGTGCCATAT AGCCTGGTGG  
21751 TATAGTTATC TATACAAATA TATTTATTAA TTGTTAATAT ATAGAGTATA  
21801 TAAATATCTA TTTATATAAT AGATATTAT ATATATTAAA TATCTATTAA  
21851 TATAATAGAT ATTTATATAT ATTAAATATA TAAAAATATA TAACATATAA  
21901 TAGATATATA TTTTATATAT TATATAAATA TATATTATA TATTTAATAT  
21951 ATTAATGATG AATTACTATA TTTGTATAGA TAACTACACC ACCAAGCTAT  
22001 ATGGTGTGTA TATATTAATA TATAATGTAT AATTCTATAT TAATATAATA  
22051 GTAACATATC AATACTTAAT ATAATATATA TTCAATTGAT TACAATCTAA  
22101 TTCAGAAAGA TTTATGTTGC CATACTCTC CTTACAATAT CGATATGTT  
22151 GTTTAAAAAT CCAGCAATTA TTTTCATAGT CTAATTTAG ATAGTTCTTG  
22201 ATTAATTAA TATGATCTCT GAAATATATC ACTGGATCTG TTGTGAATGA  
22251 TAAATCAAAA ATGAAAAATG GACATTACAT CATTAAGTTC TAGCTTGCT  
22301 TACTACTTCT TATGACATTT GATATAGAAA ATTTCTACCT TTCTGTAGCG  
22351 TTTAATTGGT GTTTCTGCA TGTATTATT CTGAAATTCT CTAATATCTG  
22401 CAAGTGGGAA TTATGTGGCT AAAATTAATA AAATGTAAGT GAAGGTAAAT  
22451 CAAATAGAA TCTTGGATT TATCCAGTTA TCTGAAAGTA CATTTCATTG  
22501 CCTTAATTCA CACTTTATAA ATTTTCTAC ATAAAGTTT TCTGTAATAT  
22551 TTGTCTTTAT AGCTGAGGAC AGATTCAAGAG AAACGCTCCA CCACAGACGA  
22601 GACCTGGGAT AGTCCTGAGG AACTGAAAAA ACCAGAAGAA GATTTGACA  
22651 GCCATGGGGA TGCTGGTGGC AAGTGGAAAGA CTGTGCCTC TGGACTTCCT  
22701 GAAGACCCCG AGAAGGCAGG GCAGAAAGCT TCCCTGTCTG TTTCACAGAC  
22751 AGGTTCTGG AGAAGAGGCA TGTCTGCCA AGGAGGGGCG CCATCTAGGC  
22801 AGAAAGCTGG AACAAAGTGCA CTCAAAACAC CCGTAGGCT TGTGTTTGC  
22851 CAGCTGTTAT GCAAAAGTGC TTTACTTTAT TGTTTCCATT CAATCTTGT  
22901 TTTCTCTAAC AATAGCATT CTAAAATACC AAATTCTTAT CCATATTAAA  
22951 CATGGAGTCA AATAGTTAAA TAGTTTTCT GTCTACGTTT CACAAACTCG  
23001 TCATAGAAGC CCAAGTAGGG CCTATATCTA GGCATTCTCT GGAAAGCCTC

Fig. 2 (cont'd 12)

18/124

23051 CTCATAAACT AGGGGTACTG GATGCCCTAC CTTGCCAGAG TTATTCAGG  
 23101 TAATGGGGAA ATAAGATTAG GTTGCTAAAG CAACAGTTAA GTTTTTTGT  
 23151 TTTTGTCTG CGTTCTTAAT GAAAGTTGG AATTTTACA CTAAATATGC  
 23201 CACTGAATTG CACTACAGAC TCTGAGAGGA ACAAGCAATG ACACTAATCA  
 23251 ATTGGAATGC TGGAGATTTG AAATATTGTC TGTGTATTAG ACTTCATGAA  
 23301 AGAAGAGAAT GAAATAGTTC TTCAAAATTG TGCCATACTT TTTTTAAAAA  
 23351 GACTCTCCCC GTATTTTAA AATAATGCCT AATTATAAAT AGTGCCACCT  
 23401 GAAGCACTAA TTAACAGGGT ACTCCAAATA TAATCATCTC ACAGATATTG  
 23451 AAATGAATTC TTTTCTAGT AATTAGCTTGT ATAGGGTTAA GTGTTACCTT  
 23501 TTTAAAAGA GTTGCAAAAT ATAAGACATT AACAAATAGC AAAACATATG  
 23551 TTTTCATTTC ATCTCTTCCA TCTCTCATAA TGTTTCTTCT GACAGCCAAA  
 23601 TTTTGTTAGC TATGCACTCA GTCCTCTCAA TATATGAGAT TTTTGATCTA  
 23651 AGCCAATACA TTTAGGAAGG GAAATAAATAT AAAGAACAT TCACATTTA  
 23701 CACATTGTTT CACGAAGTGT GGTGATATCA AACTCTACAG GCACATATAT  
 23751 TTGTGTATT CTCCTTAATT AGGGAAAACC GATGATGCCA AAGCTTCTGA  
 23801 GAAAGGAAAA GCTCCCCTAA AAGGATCATC TCTACAAAGA TCTCCTTCAG  
 23851 ATGCAGGAAA AAGCAGTGGA GATGAAGGG AAAAGCCCCC CTCAGGCATT  
 23901 GGAAGATCGA CTGCCACCAG CTCCTTGGC TTTAAGAAAC CAAGTGGAGT  
 23951 AGGGTCATCT GCCATGATCA CCAGCAGTGG AGCAACCATA ACAAGTGGCT  
 24001 CTGCAACACT GGGTAAAATT CCAAATCTG CTGCCATTGG CGGGAAAGTCA  
 24051 AATGCAGGGA GAAAAACCAG TTTGGACGGT TCACAGAATC AGGATGATGT  
 24101 TGTGCTGCAT GTTAGCTCAA AGACTACCC ACAATATCGC AGCTTGCCCC  
 24151 GCCCTTCAAA ATCCAGCACC AGTGGCATTG CTGGCCGAGG AGGCCACAGA  
 24201 TCCAGTACCA GCAGTATTGA TTCCAACGTC AGCAGCAAGT CTGCTGGGGC  
 24251 CACCACTTCG AAACGTGAGAG AACCAACTAA AATTGGGTCA GGGCGCTCGA  
 24301 GTCCTGTCAC CGTCAACCAA ACAGACAAGG AAAAGGAAAA AGTAGCAGTC  
 24351 TCAGATTTCAG AAAGTGTTC TTTGTCAGGT TCCCCAAAT CCAGCCCCAC  
 24401 CTCTGCCAGC GCCTGTGGTG CACAAGGTCT CAGGCAGCCA GGATCCAAGT  
 24451 ATCCAGATAT TGCCCTCACCC ACATTTGAA GGTAAGGATG TATAAAATGA  
 24501 TGCTGGAAAA ATATAAAGGA TAAATATGTG TTAGACACAT ACATTACATA  
 24551 TAAATGTGTG TATATATATA TTTTAAATAT GTATAAGGTA TATAATATAT  
 24601 ATATCTTACA ATTCTTAAA GTACACAGTG AGCTCTATGA AGCTTATCAT  
 24651 ATAAACAGCT AGCAAAAAAA ATAGTTCTCA TTTTGAGAAA CAGTCAAACT  
 24701 TCAAAGTTTC ACTGTCATTG TGATACTAGC AACACAAACA TCTAAGAGAC  
 24751 TTAAAAGCTG ATGGTAATAC CTAAGTGTAG TGATAAGGCA AAGTAATAGC

24801 TTGTAAAATT TCTATAGATT TCCATTCCCTC CTTTTCACAT TAAAAATTAA  
 24851 AACCAAATAG GTTTCATGA CTTTTGGCAT TCATTTCCAG TGTCATTTTC  
 24901 TTGCTGGCTC TTAATGAGTT GGTGATCATA AATGTAGATG AAGTTGTTTT  
 24951 CCTTGTAACA GATTCCATTG GACAGATTAA TACAGTGTCA TATCTTGACA  
 25001 CATTAAAGAC AATCAAGATA TGACATAATT TGAAACTATT CCAGTGTGTTG  
 25051 GTACAGTATC ACAACTGAAG AGTGGGCTAA GCTTTCTAAC TCTTCATCTG  
 25101 CTTTCTTGA CATGACTCTG GTAAGGATCA TGACTTGGTT TCTGTTCCCTG  
 25151 GATTGTTTTT GGTGTTAAAT ATGTGAAGTT CTGCTCTAAG ATATCACTGT  
 25201 TTTTAAATAC CCATGTGTTT TTAAGTGGTA GGAAAATAAA TGCAGTTAAA  
 25251 AATTGGGGAC AAATATCTAA ACCTCTCTGA GTCTGTTTTC TCATCTGCAA  
 25301 AATGGTAGAG TGTGGTTTAT AGTTCAATTAT GGGTCATAA TTTTTAATGT  
 25351 TTGTTTTAT TCTGTTGACT AAACCCAGAA CTTTGATATC TTGGAAAGGA  
 25401 AAGATTTGA AACATTTATT TTACAATAAA GCAATTTCAG ATACCTGATT  
 25451 GTTTGAAAAAA CCTAAAGGCT TTATTCCCTC GTAGTAATAT TAATGCTGCA  
 25501 GAACTGTCTT TTTAAAATAC TGATTCTCAT TGGGAAGAAT GAATTATGGC  
 25551 GTATAGGGAG AGTAAATATT TCTGTTCTT AAGTAAAGC CAATAGTGCC  
 25601 CTCCTGTGGC CCATTACCTA TGAAACAATT TCTCATATTG GTCATAAAAT  
 25651 ATTTCACTGT AGGAAATATG GATTTCATG CAACTCAATT AGTAATCATT  
 25701 ATGCCATTAC TTCATATCAT TGTATTCCA TATTACATA AATTGATTTC  
 25751 TACCATCTGC TTCATTTACA AAACTAAAAT GTTTTCTGAA CTAAACTCCA  
 25801 AAATCTAACCA GCACCAGCTC TGTTTCAAAT CACTATTAAA AAATGTATTT  
 25851 GAATAGCACT GGCAACTGAC ATAAAACCTT TTGGCCTCTG CTGGGGAAAA  
 25901 TACAGACAAA CTGACTTGTT GCCGACAATA TCAATATTGT TTCCAACCAA  
 25951 CTGCTCCCTG ACAGTGACTC AGACCACCAAG ATACTCAACA CAACTCCCTA  
 26001 AACTTGCTTT AAGCGTTCCA TCTAGATTAA GAATAAACTG TTTAAAATT  
 26051 TAAAAATAAA AAAAAAAGAG AAGAGCTCAT TTAAGTGTG TCTATCGAAT  
 26101 GCGTAGAAGT TGTTTCATTA TAATGGTTCT GTAAATAGGT AACAGCAAGT  
 26151 ATGGTCAAAC TACTGACTTT GAGTGAAAGT CTCATGATCA CTTAAATTAT  
 26201 GAAAACCAGG GGTTTTCATG TTTGACTTAC TTTTGTTCCTA CCCACTTCCC  
 26251 CTCTTCCCT AGTAGCAGCT CAGTACTGAC CTACCCCTAT ATGAGAGATT  
 26301 TTCTGCACCTT GATAAAGAAG TCCAAGCTTA TAAAAGTTCA TTAACATAGA  
 26351 GACAGGAAGT GCTTTGTAGT TCAGTACATC AAAGCACACT TGGCTCTGTG  
 26401 TACTGTAAAC CGAAATATTA AATGTGGATA TTAGCTTCTT GGAACAACTG  
 26451 AAGTTGTTAT TTGTTTTCT TTTAGGTTGT TTGGTGCCAA GGCAGGTGGC  
 26501 AAATCTGCCT CTGCACCTAA TACTGAGGGT GTGAAATCTT CCTCAGTAAT  
 26551 GCCCAGCCCT AGTACCAACAT TAGCGCGGCA AGGCAGTCTG GAGTCACCGT

09/914549

20/124

26601 CGTCCGGTAC GGGCAGCATG GGCAGTGCTG GTGGGCTAAG CGGCAGCAGC  
26651 AGCCCTCTCT TCAATAAAC CTCAGACTTA ACTACAGATG TTATAAGCTT  
26701 AAGTCACTCG TTGGCCTCCA GCCCAGCATIC GGTCACTCT TTCACATCAG  
26751 GTGGTCTCGT GTGGGCTGCC AATATGAGCA GTTCCTCTGC AGGCAGCAAG  
26801 GATACTCCGA GCTACCAGTC CATGACTAGC CTCCACACGA GCTCTGAGTC  
26851 CATTGACCTC CCCCTCAGCC ATCATGGCTC CTTGTCTGGA CTGACCACAG  
26901 GCACTCACGA GGTCCAGAGC CTGCTCATGA GAACGGTAG TGTGAGATCT  
26951 ACTCTCTCAG AAAGGTGAGC TTTCTGGAG GCATTGATAA CATCTTCCCC  
27001 CTCTTCCCTG CACTATGCCT AACCCCCACC CCATTAATT CCCTTGATT  
27051 CACTGTGAGT GCCCCGGTGC AAAAGATGT AAGACTGATG AAACCGGGCC  
27101 TTTCATTTGC TCTCATTACC AAATTTACAG AGGAATAGAA TCATTAAGG  
27151 TAGGGTGAGT GGATAATTGT GTTAATATGA ATGCATACAT TTATACCCAG  
27201 TAGGCAATGT GAATAAAATT CAAGGAATGT ATTTAGATAT TGAATGAGGT  
27251 CTCCTGAAGA CATTAAATG ATTTGGCTTA AGCTTCAGAA CAACACTAGC  
27301 TCCTTATGAT GACTTAAGCA TTTTGAAAGA CCAAATTGAA ATTATTCTAT  
27351 AGTTATGCTC AGAGCAATAT GTTAAATTG TTCCATTGAT ACTTCTATGA  
27401 AAAATAGCA GATGGATTGC TGGGAAATCC TAGTTGGCCT GGTTAAAAAA  
27451 AAAAAAAAAA TCAATTGTCA GCCATGAATC ATTAGAGAAA ATTATAGTGT  
27501 CAGTGCCATT TTCAATAGAC TGCTTAAAAA GTAATCATAT TACAAAGTGT  
27551 TTCTCATTGG CTTTATATAT ATATATAAAC TTAAAGTAGA GGACATAGCA  
27601 AGGCATTTCT TACCTAATAT GCTTACTGTG AAGCATCCCT TTTGAGCAAA  
27651 ATCACTCTAA ATTTCTCCT CAAAGTGATC CTCTCTTGAT TATACTGTAC  
27701 TGACTCTTAC CACCAGGAAA ATGTCTTAAA ACCACTTCTT TTTCTGATA  
27751 AATGCAATGC TATTTGTCTC TTGACATAAG TAAAGCTTA AACATGGTCT  
27801 TGGCCACATG TGGAAAGAAA TACTGGTCAC GTAAAATACC TGATATATCT  
27851 TTCTATGTCT TCCCCTGTTT TTTTATTTT TTTTTATTT TTATTTTTA  
27901 ACTCTGATAT TGATGATGGC ATTTATTTTC TAGACCTTCA GCCTTACTCC  
27951 CGGAATGATA TTTTAAACA TCAATTAAAG CCCTTAGCTA GACACTCTCT  
28001 GCATTACGCC AGTTTCCCCT TAATGTAGGA TGTCCCAATT TGAAATTCCC  
28051 CATTTCCTCT TGACTTTGTA AAATACAAA CCCAGAGCAA AACATTGCTT  
28101 CTTTCCCTCT TTACTTCCTA CTTGCCTAAC AATGAGACAG GGACAGCCGT  
28151 GCAAATGGGG CTTTCCGATG ATAAAGTAAT TTTAACACTA ACTAAAATAT  
28201 TGGTGTTC TATGGTGGGC TGCTAATTAC AAAATACATT TTTCCCTCCTA  
28251 AAGAAAAAAA CTGGCCAAG GCAAACAGCT CAGTGATAGC AAATAAAATG  
28301 TAACCATTTC CCTATGGTTT TGCTATTATA TGCTATTATA GACAGCATAAC

Fig. 2 (cont'd 15)

28351 GTAAAGACCA GTAAGGGTTC ATTTTTCCAC CTAATAATGTC GGGCTTCCTG  
 28401 TAAAATCTT GATTCTAGTT TCAGCACTTC TAAGGTAAAT GGGCATCTTC  
 28451 ACATGTCATT TATAAAAATT CTAATGAATG AATTATATTA AAATAGATAA  
 28501 ACAACCTATA GTTTTAATGA ATGTATCCTA GATTGTATGC TCATATGTAA  
 28551 GGATTCTAAA TATCAACTTG ATAACCAAAC CAAACATAGT GCAAATAGGT  
 28601 TATCATTAT TAACCACAAC CACCTTCCAC AAAACTGGTC ATTTTTTAAT  
 28651 TATTAAGATA ATCTGCAACA AGTTGGCCAT TTAGCCATCA GCCTATTTCT  
 28701 TCAGCATTAA GACATTAATC CCAGATTCAAG AAATAAAAGTC AAGTAACATAT  
 28751 TTATAACCAA GTAACATTCA AATCAAAACT AGATGAAAGA TTGGTTAGTT  
 28801 GCATAGCTAT AACCAAAATG CAGTTTAAT ATTTTACTCT AATCTATATT  
 28851 TTAACTGAAG TCAATAAAAT TTTCACTATG GAAATACACT AGAAAATATG  
 28901 CAATTCTTA TTCTTTTAA GCAGATTAT TTATTGTACA TGTCAGTCT  
 28951 TTGAAATAGG CCAATTATTTAT TTATGTTATG TTATGTTATT TATTGTTTT  
 29001 GAAATGGAGC CTCACTCTGT CGCTCAGGCT GGAGGGCAGT GGTGCCATCT  
 29051 CAGCTCATTG CGTCCTCTGC TACCCGAGTT CAAGCAATTC TCATGCCTCA  
 29101 GCCACCTGAG TAGCTGGGT TATAGGAGCG GACCACCATG CTGGGCTAAT  
 29151 TTTTGTATTT TTTGTAGAGA TGACGTTCA CCATGTTGGC CAGGCTGGTC  
 29201 TCGAACTCCT GACTTCAAGC GATCTACCT CCTTGGCCTC CCAAAGTGTG  
 29251 GGGATTACAG GTGTGAGCCG TGGCACCAAGC CTGAAATAGG CCAATTATTA  
 29301 AAATGGGAGT ATTCTACAT TAAAATGGCC AAATAAAAGAC TTTTCTAAA  
 29351 ATAAACTTTA AACTAATTAT GGATAAAATAT GTTTGCCTT TGAGCCTTAA  
 29401 TAAAATGCAT TAATGAATAT TAAGCTGTAA AAAGTACATG TTAACTACAT  
 29451 AGCTATAGTG TATAATATTA ATATTAATTA GTGCCTTCCA GTAAATTACT  
 29501 AGATTAAAAT AAATTTAAT ATAAGACACT GAGCTTTTG TTTCTTGAC  
 29551 AATAGAACTG CAAGCAATAG CAAATTGCTC TAATCCTTTC ACGTACATTT  
 29601 AAGAAAAGTTT ATGACCTATT GAAGAGAAAA GTAGATCTAG TGGGTGATAC  
 29651 TGGCTTCATT ATGGTTAATT AATTGATCAG TAGAATGTCA GAAATGCTAA  
 29701 GAAAACCAAA GAACTACACC AGAGAGAAAA TGTGTTAATG TAAATTTAA  
 29751 GGCAAGTTAA TTAGCGATAT ATAATAAAGA TGTATATAAG TTCATGATTT  
 29801 ACCTGTTGT CTACAATTAT AGATGATTTT TTGATACTCA TATTTAAATC  
 29851 GGTAGCTTT CCTATAGATT TTAATTTTG TTTAAATTCC TCTTCGTTAA  
 29901 ATTAAATAAA ATAATAAAAT ACACCTTTA ACAGTTTCT CTTCTGCAGC  
 29951 TGCTCTAGGT CATTGGTGGC CATTGAGCCA TAACTAGTCT ATATTTGTTT  
 30001 TGGGTTTGT TTCATGTGTC TGACTCAACT AAATTTTAA ATAATTTGTA  
 30051 GTAACCAACT TTGCAAATTC TGGGTTTGTC TTTAAATGTC AGATCTGGCA  
 30101 ACGCTGCCTT GACATTCTG CCTAGAACT ATTGGCTCTA GGCAGTCAGT

09/914549

22 / 124

30151 GTCTGTCTGC TTCAGACTGT TGACTGAAAT CCCCATTGCGT TTTCATGCC  
30201 TATCTGGCCC TTGCTGGCAT ATGAGTTTGC AACCTTGTT GATTTGCAGA  
30251 AATTGTCTAT GTTAGAAAAT CATTAATATC TAGATTCAA CATATTTCTA  
30301 AATAAAGCTT TAAATTATTA TGGTAACCTT AAATGTATTT ATTCTAATTT  
30351 TTTTCATTAA ATTGCTCTTC ATCATATAAA TATATAATTT TTATACAAC  
30401 GGATGAGTTT GGCAGAAGAA TACCAACTTT TCATATTCTT TGTGGCATT  
30451 AACTTTAAC TGTACACATG GAAATAAATA ATCCTTAAAAA TGACTTATGA  
30501 CCACATAAAAT GCCTTAGCAC ATGTGGTTCA TATTTGGAGA TTTCTCATAT  
30551 TTGTTCAATA TAATTATTT TGTGGTTA TCCACAGTAC TTAAGAAAAC  
30601 TTCTATAGTC AACATATATA CTGTAACCTGG CCTCTACACA GTATAAGCAA  
30651 TTACCTTACA TGGCTATTAC CGATAAAAGTT AAAGTTGTAT AAAGCCTTG  
30701 GATGCTTTG ATTTCAAGTGC TAAATAATGG AGTACACATA GAAGAAAACA  
30751 TTTTAGCTTT GGTTTGAGTG ATCAAATTT AGGTCAGCCT TTTTACATT  
30801 ATGTTATATC ATCCCCATTA TGCATACCT GTGTATTTAA TTTTGATCAT  
30851 TTGATGTCCT AAAGGAAGAA AGCTATAATT CTGCAATTT AATTAATTT  
30901 ACACTTTGCT TATCCACATG CCAGAGATTA TAAAAGAAAT CCCTAAACTT  
30951 GTCCCACTTA GTTGGTGTATA TCCTCTTCCT GTATTTTAG AGAGGCCATT  
31001 TCTTATTTTC TCTAGACATA GCTTTTCATT CCTTCTTGTG ACCAATTGTG  
31051 AATTCCCTAA AATAGAGATG ATAAAATTTA TAGCCTTTA AATACCTAAT  
31101 TTATGATTTTC TAAAAGATGG TATAGCTTAA TTTCATTAAA ATATTCAAAT  
31151 AAATGATACT AGAACATTAATT AAGTTTTAAG CAAACATTCA TATATCTT  
31201 TTCACATGTG TAAATGGAA ATAAACATGC CTTTTTATTA AAAATAATTT  
31251 GAAGACAAAAA GATAAGTATT AAACAAACGTT TTATACCATC TCTGTCAATT  
31301 GGAAGTTGTC ACTCTAACTT AGCCAGAGCA GATCTATCTC ATTTTGATG  
31351 TGATATCATA GCAAAAGTCT AATCAGTTGC ATAGGGAAAGG AAAAAGTAAG  
31401 ATAGTATTTA ATCAATAGGA TTCAGAGGAA AATTATGCTA ATGTGATTTA  
31451 ATCTATTTTC TAGTAATCCT ATCACTAAAC TGTCATTGAA TTGTACTGCA  
31501 TTAGAAAGGA ACTCAAATAT GTGTGACGGC AATGGACATC TTGTCACCTT  
31551 TAGTTGGCCT TTTCAATGA GTTAAGCATT ATATGTGTGT TACCAAAAAA  
31601 TTATTTTTA TAGTCAGAG AACCATTTT GTGGATGTG TAATTTGGAA  
31651 GTTTGGTTA CATTATGTCC TTAGGGGTTT TCTTGTGTT AACAGCATGC  
31701 AGCTTGACAG AAATACACTA CCCAAAAAGG GACTAAGGTA TATATTCTC  
31751 TCAGCACAAT TGCTACCTCT CTGTTGTTAT GTAAACTTTG TGTGCTGTCT  
31801 CTCTTCCTTC TTTGTTGTT TGCAATGTAG CACATGACAT TGAGGACGAA  
31851 ATCACTTTA ATTTGATGG TTTCTCTGGC CCGAACAGTT GGTGAGATAG

Fig. 2 (cont'd 17)

23/124

31901 CCCCTTAGGT AGAGATACTA GTAGAGATTG AGGCTGTCTC TCAAATTAAA  
 31951 TAAATTCCAA TGTGAATATC ACTATTTGA AGAAATAATA CTAAACAAAC  
 32001 AAACAAACAA AACAAAAACA AACAAACAAA AAACTTGTCC CAGGCATTAC  
 32051 TTTTTGGGG GCAGCAACTT TGGTAGAATG CAGAACTCAC TTCAACAAAT  
 32101 TAAAATAAAA TTAACTCTTC TAACTTTGC CTATTAGAGT CATATGCATG  
 32151 CAAATATTCA AAACCCATGC AGTCTACAGA TGTGGGCAGT TAATGTTGAT  
 32201 AGGTTGAAGG ATGCTACAAT CTGAATCAA GAAAACATAT TTTCATCATC  
 32251 ACAGGACAAA TGCTGTAATT AAGGTGTGAT TTTTATAGAA TCCTTTGAT  
 32301 AAAATCTCAA AATTGTTTA ATTTCTATTG TGCAGGGTA CTGCTATCAG  
 32351 ATCAATTAA ATCTGAATTA ATCTAATATC ATTTAATAAT CTAAAATAA  
 32401 TTATTCCATC CATAATAAAA ATAATAAA AAATTTAAT TATGCCATC  
 32451 TTTTACTGTG TACTTTATC TGAGGAAGAG ATAGAATGAT CTACTAATAG  
 32501 AGGTATAACA CTGTATGTG ATGAAAAGTT GGCTAATTTT GGTGCTAAGA  
 32551 ATTTACTTAC AAAAGAAAA AGAATATACT TAGTTGGTG AAACACTGAA  
 32601 TAATGGCGAA ACTAGGTCTT TCTCCATTAT TTTTTTCTC TCCAATTTT  
 32651 CAGCAATAGC AAATAGCTGG CAATTATTCC ATGTTAATAT TTTGATCCAG  
 32701 AAATTTATGT TCCAGTAAAG CGAGCACATC TCCCTCCTTA TTTTGTAAT  
 32751 CTAGGCATGA TGTCAAGTGG CAGTTAACAA AAAGAACTGT TTTCCCTTTA  
 32801 AAAAAAAA AAAACAAAAA GCTGCCAATA TGTATTCCAT TTCCCTATGC  
 32851 CTTCTGTGAC CATCCTTCAT TTCCCTGGC CCTGGCCAC CACTGTCTC  
 32901 CATTGTAGT CCATGTTTC ACCCTTTTA CATCCTTCT TGCCCTGTGC  
 32951 TTTTGAGTTC TCAATTAATC TGGCTGTCTG CTCATTGCTT ATGATTCCA  
 33001 ACTGCATATC TGATAGAAGC ATAATTTCT CCTCAAACCC CTTTATCTTA  
 33051 TTTTTTTCC CTATGTGATT CAAACAGATG GCGTAAGATC ATCTGGAAGA  
 33101 ACTGAGCAAT TATAATTAGA TTCAATCTGT TTGAAATTGT TCATTCTGAA  
 33151 TAGAACCTC CTCTGAATTG TTTTCTGTG CTGGCATTGC CTTGCCCTTG  
 33201 TAGATGTGCT TAAGTGTCA AGCTGTGCTG TTTTGCAGAT ATACCCATC  
 33251 ATCTCGGCAG GCCAACCAAG AAGAGGGCAA AGAGTGGTTG CGTTCTCATT  
 33301 CTACTGGAGG GCTTCAGGAC ACTGGCAACC AGTCACCTCT GGTTTCCCCT  
 33351 TCTGCCATGT CATCTTCTGC AGCTGGAAAA TACCACTTTT CTAACCTGGG  
 33401 TAAAATATTC TAAAATATTG ATTTTGTGTT GTTCTTTCA CCACCCACTC  
 33451 TCACAGAAAC CCTGGAATCT CTCCATAACA CAACACGTT TCATTAAAG  
 33501 GGAGGGATAA AAGCACTTTA ACAGTACCTT TCATTTGTGAT CATTGTTAC  
 33551 TCTTCACAGA AAAATCTCCA AACATTATGC TATTATTGC TCATGACAAA  
 33601 TGCTTAACAT AGATTAATAC TGTGGTTGTT TTCTAGTCTA GGCTCCAGAG  
 33651 GCTCAGAAAG TTCACTTGAC TTGAAAAAGT CTTACCATTA CTAAGGGTTC

09/914549

24 / 124

33701 AAGGCAGTAA CCAGTTCAGA ACATCTGACT TTAATCCCAG GGGCCTTTCC  
33751 ATTCCATTAA AGAACCTCT TAAAAAACAG GAAGGCATCT CCTTATTAT  
33801 TTGTCTGAAA TATTAAAACA TCCTTAAAC AAAATTAGTA ATCTTTGTA  
33851 GAAAATAGAA ACAATTAGGA AGAAAAAAAT ATGTAATTCC ATGACTCAAA  
33901 GTTAACCTCT TTTAACACTG TTAAAGTTAA AACTCCTTAA AATTCACTACA  
33951 AGAACATTCTG TTAAGACAAT ACTCTGAACA TTTTCAAATA GATACAATGA  
34001 AAAATAAATT ACCAACCTAG TCATTGGGTT ACTTTGTATT TAACATCATT  
34051 TGTATGAAAT ATAAAATCAT TTGCATAAAA TTTCATTAAA AGCACTCTGA  
34101 GTAACAAAAT AATTAAAGAA AACTAAACAT GCCAGATACC ATTTAATAGA  
34151 TTCAATGACT TTAAAAATAT ATTTATTTTC TATAAAGTCA CATATAAAGT  
34201 ATTTTCATTA TTTTATGGT AAATATTTT ATTATTAGTT TATCAGAAAA  
34251 ACTTGACAT AAAGATGAGT ATTGACAT AATCTTATTA GAGCCAGAGA  
34301 CGATCATTCC TTCTAGAAAA ACACATCTCT GAATTAGGA CGGAGGACAA  
34351 TGAAACAAGA AATTCACCTT TATAATTAC CTTTGTCAA CTATCCCAGA  
34401 GCACATCAAT TCCATCATGA AAGTACTCTT TTGACATTAT ATAAAAAATT  
34451 AGTAATAGAA AACACACAAT CCAAAACCTT ATATTTCTA AACTTCAAGT  
34501 TAATCATCAA CTTCTCTTAG ATTTTGAAAG ACCTGAAAAT AACACATAATT  
34551 TCAAATAACA GAACCAAAC ACCATATACA TTTGTAATGA GGCACAAACAG  
34601 TCAATTTGA GCCTGTATT TTCCAGGTT TAGCTGAATA ATCTTCACTG  
34651 CTTTCTTAGC TTTTGCCAG TCTAGTTGG GGACTATTG GCCTTACTGG  
34701 GCCTAACAG AGTGTATAT TAAAATATGT TAATAAGCCA TACTGAGAAT  
34751 AAGATAATG CAGTTTCTA ACTCCTTAGG GACACAAGTG GGGACAAACAC  
34801 ATTCCATGAA CACAGGTGAA TGAATGCCCT TAGTTCTCT GAGTTGGACA  
34851 ATTCATGCG ATCATTTTT TCTCTGAGGC CAAAGTCTCT GGTTGATCT  
34901 TCCTAGCAG TTCCAGAACAA GAAAGTGGAGT TTACTTGTC TCCATATTCT  
34951 TTTTCTCCAT GCTCGGAAT CCCCTGCTTT CCTGATCCCA CCACAAAAAC  
35001 TCCCCGAGG ATGAAGCCTT GGCTTCCAG GCTTCCAGGG AAGCCTCGAT  
35051 TCCTGGCTGG AGGTAGTTGT ACCACACTCC CAGAGGGCTA AATCCCATAA  
35101 ACATCATCTT CTGTCCTTGT AGATCATAGA ACTTTTATT ATCATCCAGG  
35151 AAGATTTCTC TTTGAAACA AGGCTGGAAA AACTTTATGT CAGTCCTGAC  
35201 CTGCTCTTAA ATGACTGCGT AGAGGGAGAT GCCCAGCTTA TCCAACTTGG  
35251 GTTGCAGAGA GGACAGATCT GCAGCCCCCTC TTGCCAGAGA AACACATCCTG  
35301 GCACAGCCAC AATCACAACT CCATTCTCT CCCGATAGCT CCTTGCTTT  
35351 GAAACTCATT GGTTACTTCT CCAGTGTGTTT CAGGTCTATA TTCTCCAGGT  
35401 ACTCCAGCAC CTCTTCCAG GGCTTGGACA AAAATACATC TGTGTTGGCC

Fig. 2 (cont'd 19)

25/124

35451 AGCATCAGTG CCAAGGCAGC AGCCTCCAAG GGCTCCTGCA CCCATGGACC  
 35501 ACATCCACAC AGAGAACAC CTTGGGTCT CAAGTGCCTC CCTCTTCTTC  
 35551 CCTTCTCCA AACCTGAAGC CCAGACACTA AGGGGTCAAA CCCTCCTGGG  
 35601 CCCTGAGGGT TCCAAGGCC TCATTACTTT TTCTTTTTT CACTGGAAAA  
 35651 AAAATTCTAA TCATGCACCT ACAGAAGATT GACATTTTC AGTAAGTTGG  
 35701 ACTTTCCAGC TTTCAGCCAG GACAAGACTC AAGGCTATGT CTTTCTATT  
 35751 GCAACCCTTC CCACTATATT GAGTAGGGCT TTTAGCAATT GAAAACAATT  
 35801 ATTTTGGTCA TGTTTCATA TAAGCTAATG ATTCATATC AAACACCAAG  
 35851 TTTTTGTTTC CTAACCTATA TAGTGATAAG AGAATTTACC TATAATGCCA  
 35901 AAGAATGTAT AGCTTTTATT TGCTTTAAGA TGCAGTTGAT TTTTTAAAAA  
 35951 AGCGAAAAGC CTAACACTTT AACTTCAAAA AATGAATTAA AAATGTTGT  
 36001 GTAGGTCTA GGAATATGAA AAAATTTAT ACAACATCTA AAACACACCC  
 36051 AAATCACCTA AAGTGCTATA AGCTTGCTAA GTACTTCATG TCTCCTATCA  
 36101 ATTCTTTCAT TAATTGACGT TAATTGATT AGTTGACTCC TTCTTCTATT  
 36151 TTTCCTCACC ATTATTATTC TGATTAATC CACCTTCATT ATTCCTTAGG  
 36201 AACAAAAAGA CTCACCACTT AACTATGTCT GACATTGGTG AAGTCGTTA  
 36251 AACTTAATTT TCTTATCTCT TGAATGGATA CATAATACCT AGGTTATATT  
 36301 GTAAAGAATG ACGGATATAG TGTATGTAAA GATGGAGAAG TGTGTAAGAC  
 36351 TTGACAGATT CTGCCAAATC ATTATTTCA CTGGAAAGCA TGTCTTACAC  
 36401 GATCATAGAG TAGCATTCTA CAGATATGCC TGAGCTTGT CTACATTTAA  
 36451 TTGAGTAGTA ATTGCAACA CAGTAACCAC AGGATTTAT GTAAAAGACA  
 36501 TTCACAGATT GTGTTTTGA AAGATTGTAT TTTGAAGTA CAAAACATG  
 36551 ACATTGTTAT CAAGGACTCA TTTACCACAA ATATCAAATA TTTGTGCAAA  
 36601 GATAAGTTA TGCTAAGATT TGCATAAATT AAAGTTAACCA TGCAACTGA  
 36651 AGCTAACATG TCCATGGTCA CAATGTGTTA AAAATGAAT GGTTCTGTAG  
 36701 CACACTTGGG AATGTATTTT ATTACATAGT TTTCAGAGTT AAAACACAAT  
 36751 TAATAAATGA AATGTGAATT ATACTTTAC TGACAACAAA GCTCTCTGTA  
 36801 GAGCTTTAAT GTCTAATGA ATTAGAAAAC CACTGATCAA ATACATCCCT  
 36851 TACATTTCAT TGCTATAGAA ACCAAGTCTG AAAGGTTAAG TTTACCTTTC  
 36901 TAGGATGTGG GTTTCCCCCCC TTAATCTATT GTGGTTTATA TCAGAGATCT  
 36951 CTCAGCTGTG TCAGACAGGC CATGACTTAA GTGACACTGC CCTCTTGATT  
 37001 CTCTTCATAC TTTTCCAATC ACAATTCTTT CTCCTGGGT TGCTCATCTT  
 37051 AACATAGCTG TATCATTAT TGTAGACACA AGGTCACTTT TGAGAGTGAA  
 37101 TGGGACTATA TTAATAATTG TTCCAGGTAT TAGGTGCAAA CCCTGGGCAA  
 37151 TGCAATTCTA CCTCCATCTC CTCCCTATAT TTATGTGTTT ACCAAGTTGT  
 37201 TTTTCCTGTA GACTTTTTT TATCCTAAAC CCTTTTTCTA TGTTCTCATT

26/124

37251 CACAACCTTA ATTCTAATCT CTCAAATCAA CATTTCACTT TCTGTCTGAG  
 37301 ACCTTTTCA GCTCTAAAAC TAAAATCCCA TCAGTGTGCT AGACCATA  
 37351 GCCACCTGAA ATCAAAGTCT TTTCTTAAGT TCTTTCTTC TATTTGTCTT  
 37401 ATAATTCAT GTATCATCCT TCTCTCTACT CTAGCACAAA ATCTGTGTAA  
 37451 TCAATAGTCT TACTTGAAAC TGTGCTCTTC ATATTGTACA TTTTCAATAG  
 37501 ACAGGAACCT GTGATTTAT CTTCAGAATA TCTCCTACAT CTGTCTCTCA  
 37551 TTTTCAGGGA CATTGTCCTT GCTGAAGCTT TTTTAACAT AGACAATTGC  
 37601 AGCAGATTT AAACTGATCT TACTCTGTCG ACTCCCTTAT GTTTCAACAT  
 37651 TTTCACCCAT TGGAAGGTAT AAAAGAAGAT ATTCCTGTCC GTGTCAACAT  
 37701 AATCTCATGT ACCTCTCCAG ATCTTAGAAA CACGTATGGC TTCAAATCAG  
 37751 GCATTTGGAG ATCTTTATGC TGTATGGTT CAGAGTGGAA AAAATGATTG  
 37801 ATTCAAAAAC ATAATATTTA AAGAGTTTT ATTGTATTTA CAGTCACCT  
 37851 GAACCTCTGT TCATTGGCA AGAAAATGAG TACTCTTAAA ATGCAATAAT  
 37901 AAATTAAAGT TACTTTATTA TTAAATTTA AATATATATA TATATACTTA  
 37951 CCTTAAATAT GTCCTCTTGT TGTCTTTAG CATCACCCAT TTTTGATTG  
 38001 ACCATTATCT TTTCTGAATA ATCAGTAAGA TACAGGATTA TTATTAATGT  
 38051 TCAAAAGTTG CAGTATTCTAT GTTTCTTTA TTCTTCTAC CAATTAAAAT  
 38101 GTGTTAATAT ATAAAATTT TAGAAATTT ACTAAAAAA ATCACAACAT  
 38151 ATATTAGAAA ATTAAGATCA CTACAATATG TCATATTAG TAGACTACTG  
 38201 TGAGCTACTG CCACAGTAAA CTATGGTCG TGTGCGTTC CCAGCATGCT  
 38251 AGCCCTAGTA GAAACCATTG CCATTCAAGA AAGACTAAC AAGTATAGCT  
 38301 TACATAAAC TAAAAGTCTT TGGATGAAAC TTCATTTGGG AAAATAACCC  
 38351 AATCGCTACC CTTCAATTTC TTATGAATGA AAAATGGAA GAATAAAGGC  
 38401 CTCTAAGATC CATTCAAAGC CAGGAGACAC ACAAGAATT CTAAATAGAA  
 38451 GAGAAACAGA AGAGGTCATA GTTCTTGTA GCCATCTCAT AACCTGGTGA  
 38501 GACTCATTGT CATGCCCTCCA TGCATGATAA CAATCGCTCA GATTCAATT  
 38551 TCATCTTGCC ACAAGGGTTA CATGCAGGAA CATTAAATGTC AACCTGTCAC  
 38601 TTCTAATATC CATCTAATAT TCTCTAAATT CGATGGATCC TTTTGCA  
 38651 GGTGATTGTT AAACACCTTT GCATAGGAAC AGTTTCTATG CTTTTGTACT  
 38701 CAAATCTTCC TCTACCTTGA ATCCTTCCC ATCTCGTGT TCAACCTTCA  
 38751 ATCTTCTCAG AATGAACCTCC TGTCTTCTAT TCTTTCGGAA GCATAGAAC  
 38801 TCACGGTCAG AAGAGACCAC ATCTGGTTCA ACCCTTCATC TCTTATGTAA  
 38851 AATTTTATGA CATCTCTAGC TTCTTCTTTA AACCCACCAA TGACAGAAC  
 38901 TACTAAAATC TAGAAATAAC ACCTTTGAAA TTCTTCTTT AAGAGATCAA  
 38951 ATAAAATTT CCTGAATCTT CACCTATTGT TCCTAGTTAT ATATATCCAG

27/124

39001 ATTCTACAAA ATAAGTCAAA GTTAGATTGC ATATGACAGC TCTTCATATT  
 39051 TAAAACAATA TAATAAACTC ACTAGTTAAT GTCTAGCTGT AGATGCAAAA  
 39101 GTAGAGAGTG ACTTGGGGTT ATTTAAAAAC CCAGTCCAGC CAGACACATT  
 39151 GGATCATGCC TGTAATACCA GCAGCACTCA GGAGGCTGGG GCAAGAGGAT  
 39201 CCCTTGTCCA GGAGTTACAG GCTACAGTGA GCTATGATCG TGGCACTGCA  
 39251 TACTCCAGCC TGGAAAGACAG AGTGAGACCC TGTCTCACAA TAATAGTATT  
 39301 TAATAATATC ATAAAAAACCC AGTCCACATT TATATAGGAT CCTGTTTCC  
 39351 TCAAGTTACT ACAAAATAAT ATATAATCTT AATAAAAGGT TAGTGGCTTT  
 39401 GCCAAGATAG TGGCTTGGCT ATGCAAATGC AATTAAAGAC AAAGTTGGTA  
 39451 GCCCTCTTT TCCTAATACA TTGCCATATC TGTTCTCTT CTATTTGGAA  
 39501 ATTCTTGTGT GTCTCTTGGC TTCGAATGGA TCTTATAGTC CTTTTATTCT  
 39551 TCCATTTTT AGTCATAAAA AAACTGAAGG GTAGTGATTG GGTTATTG  
 39601 CCAAAGCAGA TGGAAAGCAA AACTACCCT AGAACGCTTT TACCAATTG  
 39651 TGTTCCATTC AAAAAATTAT CTTGTATGT CTTACATTTG TCTTCTACTG  
 39701 TATAGTTTTT CTTGTTCTAT TTTACATATT AACTTTCTC CTTCTTCAGA  
 39751 CATCTGCCCT ACTGGCTACT CTTGAAATCA GAGACTGTGT CATATTTTC  
 39801 CTTCTATTCA ACTACAACAT CTAAAAGCAG ATCTGTCATA GTTATTAAC  
 39851 TAATTGAACA CTCTAAATA GTTAGGTGTA ATTTCCAATG CAGAAGCTAT  
 39901 CAAAAGGGTT TGTAAATGCA AACTATTCCC TTTAAAATCT ATCCTAATCC  
 39951 TCATTAATGT TTCATCTTGA TAGAGCTAAG TATTATGTAT TGAAATTGTA  
 40001 GAAGTACACT TCACCTGGAT ATCTCTGCAA TCATTTAGGT AAGAATTATA  
 40051 CAAAGCCAAA AAGCAAATAA AATATCCTCC TAACCCCTATA GATACGTATA  
 40101 CTAAAATGAT GCACCTGCAA ATTTGTTAA TACTTCATTA ATTTAAACAA  
 40151 GAGTAAATTC ATACTGTGAA CCAAGAATAG GGTGACTTAC CCCAATCTTG  
 40201 CCACCTAAA CATAAACATT TTAAGTCTTC AATGTCCTAC AGTGTACCTA  
 40251 CTGGCTGTTG TCACTAATCA GACCGAAATG GTACTAATGG TCACTGCAGG  
 40301 CTGAAGGAAT ATGCTTGAAA GATAGGCAGA TCCTCTCCCT CTCCCTTTT  
 40351 TAČTTTTTC GCCTTCCAT CCTTCTTCT TTTTTTCCAA TAGATTGTGC  
 40401 ACTTTGGAGA TTCATATTTC CTTCTTTTC CATTACATTT TAAATATGTG  
 40451 ATTCTTAGTC CTATGCTTCC TTTTACTCCA ATCAATAACT GGCTCTATCA  
 40501 GAGGGTTGTT CTGTGTGTTA ATTCGGTTAA TACCAGGATT ATCAAGCACA  
 40551 GTGCCTTCCA AATGTGAGAT ACTTCTCTCC GGTTACCTCT GGGTTTACTT  
 40601 TTCCTGTTTT ACATTGTTTT GAGAGCCAGT ACTTGTTAAGA AGAAGAAGTT  
 40651 TAGTGCCTGT GTCACAGAAA AAATCTTAGT AAATTTGAA GTGATGTCAG  
 40701 AACAACTCTA AGCCACTGAC GGATTCCACA GGGTTTGAA AATACTCGTT  
 40751 AGTTCCCTTT ATATCTTAAG AGGCTCCTGC CTGCTTTCTC ATATACCAGT

Fig. 2 (cont'd 22)

09/914549

28 / 124

40801 AACAAACTTG CTTTCTTAA ATATGAGCAT TTAGAATATC TTTCTCAATT  
40851 TTTCTGTTT GCTTTTATTC CAAATTCAC AACTATATTG TTTTCCAATG  
40901 TAGTTGTACA TACATCAAC CAAATCTTC CTTAAATTGA TGACTACCAG  
40951 GTGAGGACTC TTTGGCAATA AGCAATAAGA AAATAAATTG TTATTAAAAAA  
41001 TTACAGACTT AAGATACTTC TTTGGAAATA TAACATGTTT GTGACTTTG  
41051 ACCATCTCAT CATGATATGC TCATCTTAA CAGAGTAGAA AATCATTCA  
41101 TATAATTAAC TTTATGGTGG GCTGCAGATA CCATGTATGT TACATTGTGT  
41151 TTAGTTATAA AAATGTTAT TATACACTAT TTCCTTATAA TCTAACCTTG  
41201 ATAATAATGA TGGCCTAAT CATGAACCTA CATCAATTAA GAGCTTGAAG  
41251 TGACTGAGAG TATTGCGCTG GAAGCATTAA AGCCCTTCT TGGGAAATT  
41301 AGATGTTTA TATTTTACTT TCTTTTGAT TTTGCTTTT CCATTAAAGT  
41351 GATTACTATT TTTAAAGAGA AAACCGAAAA CTCTAGAAAG ACCATCTTT  
41401 CTTCATAACA GGTAGCAGAA AACACCAGT TATTACATT TCTAGCAAGAG  
41451 CAGTAGAGGT GACTGTTGG TTTTGTGTAC TGTTGCTTTA GAAATTGATG  
41501 TAAGGCTTCC CATAAACGTG CCAGAGGAAA AGAGGGACGC AATGGGATCT  
41551 GTTATTGAAC ATTCAGAGG CAGACTCTTA CCTTAAATAG GGACTCACTA  
41601 TACATTCAAG TTTTCATAAG TATTGGGATC ATGTTCTTAC TTTCTATCAA  
41651 CCTGCTATT TCATCTTCA AGCTTAAGAG TAATAGGCTC TGTGTGTTT  
41701 GTTTTCAGT GAGCCAACA AATTGCTC AATTAAACCT TCCCAGGGCCC  
41751 AGCATGATGC GCTCAAACAG CATCCCAGCC CAAGACTCTT CCTTCGATCT  
41801 CTATGATGAC TCCCAGCTTT GTGGGAGTGC CACTTCTCTG GAGGAAAGAC  
41851 CTCGTGCCAT CAGTCATTG GGCTCATTCA GAGACAGCAT GGAAGAAGGT  
41901 AAGCGTTGAG GGGGATTAAA GATGAAGTCA CTTTATTAA ACCCTGAGAG  
41951 GGAAACCATC GTGTCACTCA CATCACAAAG ATTCTGAAG AGGAAATAA  
42001 ACTAGTGTAA TTATCATTG GGAAACTAGA AGCTTGAAGA AGTTTATT  
42051 TGTATTATCT TCTATTCTT TATGTATTG GAAATATGCC AGAATTGTT  
42101 TATATTAATA CTTGGCTGTA GAAGAGTTA GACTAAATCT ACTTTCCAA  
42151 TACAGAAATA TACATATAAA CTATTTCCC AGGTGCATCA AATATCAGAG  
42201 CAAATGTTT GTTGACATT TTGGTTAAAG AGCCATAAG ACACACAAAC  
42251 CAGAAACATT ATTTATGAA AATACCACAT GTTGCTGACT TTTATTCCCA  
42301 GGAATTCCCT CTGGTGCTAA TTTTTATTAA TATCATTAA GAATTCAAT  
42351 TGTACCTACT TTTTGCTTT ATAAGTCACT ATTTCTTCAT CCAATGGCAA  
42401 TAAAATTGTC ACCTAACCTA ATAAATATCT TTATAGTTAT ATAGTTCTAT  
42451 GTAAATACTC CAAATAAATC AGCTTGAAAA CCTCAGGAAG CTGAGTTGAT  
42501 GCTCAAATAT ATATATTGTAA GTAAACTGTA GAAGCTCAA TGTCAAATTT

Fig. 2 (cont'd 23)

29/124

42551 ACAATAATT TGAGAGACTT TTCTCTTG A TTTAATGAAT TTTTTTAGTA  
 42601 TCCATAAAGA AAACTTACAG CATACATATT ATAAAGCATG TCAGCTAAGG  
 42651 ATAAAATAAA ACTAGACATA CAAATTCAA CTGATTAGAA TGAAATTATT  
 42701 AACCTAATA ATTATGTTA AAAGAAAAGT CTCCAAATCT TGAGACATAC  
 42751 CAGAGTTAA GTCTCAGCC ATCCATTAC TTGTGGTATA AACTTAGGCA  
 42801 AGTTTCTTAA CCTTCTTATC CCTAAGTTCT GCATCTGTAA CTTCTTAGGT  
 42851 TTGTCACAAG GATGAAATAT GAGAACAAAG AATAATTCTG TTCCATGATC  
 42901 TTTTCCCTTC CTACCTTCTT ATTTAAAGTA TCTTCTGACT GAGGGGTTAG  
 42951 GCAGCAATGA AAATTGACTC ATGTTTTCA GGTCACCACT ATGGATTCAA  
 43001 TATACTGGCA TTAAATCAGT AGAGAATAGT TGTCATTGCC TTTTGCAATA  
 43051 TTAACCAAAC CACTCAGTTC ACTGTGACAG ACAGTGAATT ATATCCAATG  
 43101 ACTCCACTGA TTTTTCCAT GTAGATAGAC AAAATATAAC TACTCTCAA  
 43151 TGTAAGGACC CTGCTTCTG AAATGGTTCT GTTGCTCT TCACAGATAG  
 43201 GCTTCTTATA ATACTTTAA AATAATTGC TAAGCATAAC GATGGCTTTC  
 43251 TAGAGTGTGG CATTGACAAA TAAAGTGATT TTTATATACT GGGAAATTCT  
 43301 GGCCTTCAAT GTATCAGGAT TAAATAATCT GAATTCTGA AAGCTAGCCT  
 43351 AAGTGGCAA GATGGCTTT TTGTGCTCAC GCATTGAATA CTGAACATT  
 43401 CTAGTTCTTA AATGGCGATC TAGATTCAAG ACTTATTGAA CTAGATTGAA  
 43451 GGGACTTTAT TGATATCCTA CCTAATGCTC ACACGTGACAG ATGAAGAGAC  
 43501 TGAGCCACAT GTTCTAAGGT CATAAACAGA AAGAATGAGA ATGAGATGGT  
 43551 CTAATTAATT GTCCACCTTT CCTATGGTAC ATCAGGGTAA CACTTTAGTT  
 43601 TACGAGGGTA TTATTAGAGA TAGAAAGAAT TTTTTTTAA ATAATTGACT  
 43651 CAAATACCAA CATTTCGAC ATTACATAGA GTAATAGCTT TGCCCAAGTT  
 43701 AGAAAATGG GGGTTCTTCT TTATTCCCT TTTGACCACA TCTATATACT  
 43751 CAGTTTAAA AAGGTTCTTC CTGGTATCCT TCAATTCCAT CCCCATGTTT  
 43801 TCATCTACAA GCCTAGTGCA GCTATTCCAG CCGTCTCCTG ATCAGGTCTT  
 43851 AAGCACCTCC CATATGTCCT TGTAGTACCC ACCATATTGA TCTCAGTAGC  
 43901 AATCACAGTA CTCTATTGTA AATATCTTT AAATTATTAT CTTCTTTG  
 43951 AGCTTTGGG ATTTTATCTT ATTTATTGTT GTAGTTCCAG GATCTAGCAA  
 44001 CAGCTTGTCAT CATCGTTCACTAA TGTTGTTA ATGCACAATG  
 44051 AGCAGAAATA AACATACTAC TCCATAGTAA AAAGAGGATG AACTTTCTG  
 44101 CAAATATTAAC TCAGCACCAC TTTATCCACC TTTTGGGTTT AGTACATTGG  
 44151 AAGTATAGGA GTATAAAGCA GAATGTCAA TGTTACAGT GATATTGAA  
 44201 AATAGATAAA AGCCAGTGCG ACATTCCAT TCTCAATTTC TCTGAGACAT  
 44251 CA~~C~~CTTGAAA AAAAAAGTA TTTTCTCTT CCTAAATTA GTAAAGGAAC  
 44301 AGTAATTCCA CATTATAAG AGTATGATCA ACGCATCACA GATAATGTTG

30 / 124

44351 TAATAACACA TTAGATAAAA GTGCTTATT TCCTGAAATT ATATGGAGAA  
 44401 AAAAATCTGA AAGTGGACCT TTGTTGGATA CAAATGAAAT AAATAAGGTA  
 44451 CATAACATTT TTAAGGTTCG AAAGTTTATG GCAACTTTAG TTTGGGTTTC  
 44501 CATGCTATTC TATTTATTAT ATGGGAATT ACTGTAGCTT TCAACATGTA  
 44551 CGAACACAGGC TGGTAGGGCT CATGCTTGTA GGCTTCTGTC TAATAACTTG  
 44601 GCAACTGAGG TACTTTAGGG AGTATGGATG GGGCTCTTCC ATGTCTAAC  
 44651 GTCCTGACTG CCAAAAAATT ATAGCAGGCT GGTTCTCAGA ATCTTATAGT  
 44701 TAGTTGTTAT TACTTAATT CCCTAACAC CCGTTCTTTA CTTTTTCTGT  
 44751 AAAGGCTGGA ATTTTGAGT AGACCTTATT GTTTAACTC TATTGTTCTG  
 44801 TTTGTTTCT CCAGTTCATG GCTCTTCATT ATCACTGGTG TCCAGCACTT  
 44851 CTTCTCTTTA CTCTACAGTA AGTAATGGCT GTTAAGAAAA AGCTTGTGCT  
 44901 TTTGCCATGC ACACAGATGA TGAAATAGAT CATTAACTG TGAACAGATC  
 44951 ACATTCATCT ATGACTTGCA CAGGAGTTGT GTAGCAAAAT AACGGCATAAC  
 45001 TCTAAGCTGC CCAATACCCA ATAAAGTGCC AGGTGCTCCA CCTGCCATT  
 45051 TTTGGTCACT TACATGTGCT TTCACTTGGC TTTTGTGCAC TCATCATAAT  
 45101 CAATGAGTGG ATGTAGAATT CGATTTCATATA AACCTACTG AGGTATGACT  
 45151 TGGAGTCTCT GAAACCATGT ATGTAGTCTG CTATACTATC ATTTTAGTAA  
 45201 TGACGAGTTG TCCATGTTT GTTCTTGAG CCGTGACTGT TAATTGTTCT  
 45251 ATAGTATTTT CTTCTCATT TTTATTTTA AGTTTATTGT TGAGAGGATT  
 45301 ATCGAAGGGT AAAAGCAGTA AGGGTAAAGG GTAAAAGCAT AAAAGAACCA  
 45351 GAGATGTTT TTTTAAATA TACCTTTGA AAGAGTGTGA TTTTTTAAC  
 45401 TTTTATTTT ATTTTATTTT ATTTTATTAT TTATTTATTT TTGAGTCGAG  
 45451 GTCTTGCTTT GTCACCCAGG CTGGAGTACA ATGACACAAT CATACTCAC  
 45501 TGCAACCTTG AACTCCTGGG CTCAAGTTAT CCTTCTGCCT CAGCCTGTCA  
 45551 AGCAGCTAGG ACTACAGGCA CGCACCAACCA TGCCCAGCTA ATTTTTAAAT  
 45601 TGTTTAGAG ACAAGGTCAT TGCTATATTG ACCAGACTGA TCAATACCCA  
 45651 TGGCTTCAAG CAATCCTCC TGCTTAGGC TCCCCAAGTG CTGGGATTAC  
 45701 AGGTGTAAGC CAGCACACTT AGATAGAAAC TTTATTTATT AAGAGAAAAA  
 45751 TACCAAGTGT TCAAGTTCTT TTGCAAACGT GTGACATTAT AATTCAATT  
 45801 TGACAAGGAG AGTTTTCTG TTTGGTAAAT ACAATTCTAT CTTTTTTAAA  
 45851 AAAGTAGCCT ACAGGAAGTT ATATTTATG AGTGAGTCTT TTTAGAGCTA  
 45901 GGTAAACAGT GAGGTATATT TAAAAGCAGC CTACTGAATC TCAATGGGAC  
 45951 TTGAGTACTA TGAATAAGCC TTAATCCTGT ACTGTAAGGT TCATGAAGAG  
 46001 TTCATAGCCT CTGCTGTCAC TGATCAAATG AGCATCATGG GCAGTATT  
 46051 TTTCACTCAT TATCATTAGG TTCAAATGTT TGTTGAACC TTCTCTTTAT

09/914549

31/124

46101 AGATTAATCT CATATATTCA CTGCCCTTACA TAGTCATTCA AAATCTGACT  
46151 GTTATTGGCA GAAGTAATAT TTTTCTAATC TCTCCCTTCA ATGATTAAAA  
46201 TTACCCATAG CTTCTAGAAA TTAAGAAATC ACGATTAGTT TTTAGGTAAA  
46251 TGTACTTTT GTGCAAATGG ATAAAGTGAG GAATGTGTAA ACACACATGA  
46301 AAAAAACACA TAAAAGAAAT ATATTAAGAC TTAGTGTCC TCCTGTTGGG  
46351 CCAGCACTGC CATTGTTGG GGAATTGTAT TCTGATTAA ACCATTGCCA  
46401 TTTACATCTA TGTGTAACAT CAAAAGATGT AGCATCATTAA TTATTCTAAA  
46451 TACATACAAT AATTAATATT TGGATAAACG TACCTTCATG AACCTAACAGA  
46501 AAAACTAAAT TAAAAAGAAA GAAAGAAAGA AAAATACACT TAGATAGAAG  
46551 AAATAAGGTC TAGTGATTGG TAGCACAATA GAGTGACTAT AGTTAACAAAT  
46601 AATTTATTGT ACATTTCAAA ATAGCTAGAA AAGAAGATTT GGAATGTTCC  
46651 TAACAGGAAG AAATGATATT CTTCCTAAAT GAAGAATGGG ATATTCCACT  
46701 TCCCCAGATT TGATCGTTAC ACAGCATATG TTTGTATAAT ACCACATGCA  
46751 CCCCATAAAAT ACATACAAC TTTGTGTATC CCAATATTAA AGATTTTTT  
46801 GAAAAATTAA TTCCTCAAGA AAAGGATCAT GAGTTAACAGA AAAACAGAT  
46851 TACTAGTCTA CCAGTGTCCA GTAGACCTTT CTGTGTTAAT AAAAGTGTTC  
46901 TGTATCTACA CTATCTAATA TAGTAACAT GAACCATAATG TTGCCATTGA  
46951 TTATTTGAAG TATATCTGGC AAAGAGATGA ATTGACTTTT TTATTTTAAT  
47001 TAATTTACAT TGAAATAGCC ACATGTGCCT AGCAGCTACT AGATTGGATA  
47051 GTGCAAGTTT ATAGAGAACAA CAAGGGGTAC ATTTGTAGAT AGGAGTGGGA  
47101 TGTCAAAATG ATGAGGATAA TTAGAAAGCA TACATGAGAA ATATTGTTT  
47151 AAGAGTAGAA TATGAAATGG GAACACAGAT TAAAATAGAG TATGTATATA  
47201 TATACATATA TATGTGTATA TATATACATA TGTATGTGTAA TATATATACA  
47251 TATATATGTG TGTGTGTATA TATATATATT TATAGGCCAA TATATGGAGG  
47301 TAGGGTATAT CCTAGTGTAA AGTGAGTAA GAATGGATTA GGTGATCGAG  
47351 CCACATGAGA AGGTGATATT ATTAGAAAAT TGAAAGTTGT ATTTGAGATG  
47401 ATGAAAATGA TATATTTGAA TTGAAAAGTA AACTGTAGTA AAATAATTCA  
47451 AATAAATGAA TATTTGGGAA ACTACTTAAG AGAAAATCA TAAAACATGA  
47501 GGAGTCATTC TTTCCCCAGT CCGCCATGAT CAGGCCCTTAG GATTTAATTG  
47551 GCAATGAGAA AATACCTATG AAAATGCTTT TTAAACTATC ACATGAAAAA  
47601 GCAATTATT ATTTCATG CCTTCTTAAT AACTCTCAAT AGAGATTAG  
47651 TTGATTTGCA TTTTGCCCTG GTTCAATCAA GAAATTATCG CGTGACATCA  
47701 GGCAAGTGTGC CAAATTCTT TGGACTATAC CTATAAAATA AAATTGAAA  
47751 ATATTAGCTA GATCTAACCC ATTTGTCTCC GGATGTCTGC AAAGTGGTTG  
47801 GAAATCACAA GCCTAACCTG ATCTGCAGAG GTGTTACCTT TGGCAAACCTT  
47851 ATGGTTTTTG TGTTGTTTT GAAATCTAAG GCCAAGCGCG GTGGCTCATG

Fig. 2 (cont'd 26)

32/124

47901 CCGGTAATCT CAACACTTTG GGAGGGCTGAG GCGGGTGGAT CACTTGAGGT  
 47951 CAGGAGTTCG AGACCAGCCT GGCCAACATG GCAAAACCCC GTCTCTACTA  
 48001 AAAATACAGA AATTAGCCCG GTGTAGTGGC ATACGTCTGT AATCCCAGCT  
 48051 ATTTGGGAGG CTGAGGCAGG AGAATCGCCT GAACCTGGGA GGCTGAGGCT  
 48101 GCTGCAGTGA GCGCCACTGC ACTCCAGCCT GGGCGACAAA GCCAAACACT  
 48151 GTCTCAGAAA AAAAAAAA AAAAGGAAA GAGGGAGAGG GGAGGGAGAG  
 48201 GGAGAGGGAA TCTAAGCAA CACTGTGAAA TATTGTGAAA TATGGAGCTT  
 48251 CTACCTAAAA ATTCAAAATT TTAAATTCTT TTTAAAATA ATTGGAATAT  
 48301 CTATGGAATA TCTAGCAATA CTAAGATGAA ATTCCTCTGG GTTTTCAGTC  
 48351 ACCTGTAATT GACACCTTTA GATGTTGGCA TGGGCTCTCA GGAAGCCACA  
 48401 GCCTCCACCA ATGCTTTCT TCCTGACACT GAAGCTAAAT TTGGGTGGCT  
 48451 AGTTTCATT GTGCTGTTGC TTTCCTCATG GGAAAGAAAT ACCCTTGCT  
 48501 ATTTATATTG CTGTCAAATG GGAAATGAA AGACAGCAA GGAAGATCAT  
 48551 GTGACTATT AAATACTTCA AGTCCATTAA TTCTTTATTA GCCTGTCCT  
 48601 GTTAGGCATT TAAATTTTG ATCCCTGCAA TAGATGTTT TTGATTAAC  
 48651 GTATATTAAA AACTATATT AACCTGTTT GAATTGAAAT TCTAAATTGT  
 48701 ATTTTTCAT GAGAGCAAGT GTCATTTTG ATTCAATTGT GATTGTTAA  
 48751 CATGTTGCCT AACAAATAGC TAATACTAAC GTCATAACTT TTTAATTAGT  
 48801 AAATTGAAAT GGATAAAATGG CCACCTTATTG GCTTATAGAA TAAATAAAA  
 48851 CATTTCATT CAGTCAAGTG TTCATATTT TTTATCATCT CCAGGACATT  
 48901 GGGCTTGCTC AAAACCATTG TTAAAAAAAAA AATGGCAAAT AATCCAGTTC  
 48951 CATCATGATA TCATTAATCC CACACCTAAG CTACTGAAA AAATATATTA  
 49001 ATATTCTGGC TCATTGCTTT ATTTTATGG TAACACCCAC CTGGTATTAA  
 49051 TAACCACAGA GTACGAAAGA AGGCAAAGGT TAAAGCAAAT AATAGTTTG  
 49101 AAAAATTGGT AGTAAAAAAA GTCATGCTAT ACGGTATGTA TATAATAGAT  
 49151 ATTTAATGAT TATGCTTGCT ACTAGTATAT GTAACAGGAC TATTATAGAT  
 49201 TAACAAAAAT GCGGTGAGTA TATTCTTGA TTATTTTTTA AAAGAATAAA  
 49251 TTATTATTAA AAAATACATG AATTATTAT TGATTCTTGA ATCTTTACCA  
 49301 GCTTTCTATA ATTCTAGGAA GCCTAGAAC AGAATTGGGC AGGATAAACT  
 49351 GGCAAAAAAT GTAAAAAGTA GGCGGGCAC GGTGGGCTAC AGTGAGTCGT  
 49401 GAATGCGCAG TGCACCTGAG TGATAGATCA AGATCCTGTC TCAAAAAAAA  
 49451 AAAAAAAA AAAAGAAAGA AAGAAAGAAA AACAAACAACA AAAACAAAAG  
 49501 CAAAGTACTA GGGAAAACTA ATAGACATAG TTACATAGTT AATTGTGCCA  
 49551 TATGTTTAA GGCAATGAAA CTTTATCTT AATATTCTT GCTTACTTTT  
 49601 TATTCAAAAAA CCAAACGTG TATAAAACCT TAAAATTATT AGGATCTAAA

33/124

49651 AAATAAAATC TTTCCTTAAA AATCTAAAAT TGAGATGTAA ATTATTCAAG  
 49701 AGTGCTTTTT AAAACAGTTT TCTTATAAAG GCTATTAGGA TTCTACCACT  
 49751 TAGCCACTTT ATTATTTAGC CACTATATTA CTAAGTTAC ATATTTTAA  
 49801 AGGTAGTGAA AATATAGGGA AGACAAAGCT CAGGTTAAAA GAGTTTCTGG  
 49851 CAAATAAAAT ATATCCTGAT GGTTAGACTA CTTTGCTTA TGTTTCTGA  
 49901 AAGAAAAGCA GTAAAAAACA GTTCAGGTAG TTTTGTTCA ATTAATCTAG  
 49951 AACTATACCA AAAGTAGACA TAGAAAACGA GAGATTGTTT TTCAGCTTG  
 50001 GATCTGCTTA TGGCAATAAG CAGACTTGT A CTATTCAACA ACATTATGCA  
 50051 TTCTTCAACT TTTCCCAGAA TAAGGGAGCT TCCCAAATGC AATGGTGCAC  
 50101 ATAACTCATT TTCTGGCATT TTGCAGCCC GCATGAAGAA GAAAAACAGA  
 50151 GCTAGGAGTT TTCTGGAAGT CAAGTCAAA ACACCCCTGCA AATTCTATG  
 50201 GCAGTCCTCC TTTCCATAAG CTGCATAGCC AAAATGTTT GCCAGACACT  
 50251 TTTATCACTG GGTGTTTCAG TGTTTTCATT GTTTAAGCGT TTTGCTGACT  
 50301 TGTGATAATT AAAATTATTA ATAATCATTA AAGAAAGAAA AAGTAGAAGT  
 50351 AAATAATGTT AATTATCTGT GGTTATCAGT AGAGGTCTGT ATGTTACCC  
 50401 AGCTTTATTT GACATTGTTT GTGATCAGTA AATCACAGAA TAAAATTCTG  
 50451 ACATCTAAC CTTGGCTAGA GGTCTCTATA ATTTTATGGA GTCTGTTCC  
 50501 TACAATCTGT ATGAAAGATA CTTCAATATT TTAAGTTAC ATGCACCCAT  
 50551 CTTTTTTAGA GTATAATTAA ATAACTATTT GGTTTATGTT GCTTATGATT  
 50601 TACATCTTAG AGTCTTTAA TTCTGTCTT TGCTTAAAGG AATATTATGG  
 50651 ATCAAATGAC CTATATTTA AGAATACCTT ATGGTTATA TATTAAGAAA  
 50701 CATTATATA AAATTCTAAA GTAACCTGCT TGTACTATTT CAATTGAATA  
 50751 ACTTAATGTA TTTCATTCTA TTCTTCTCAT AGTAGATAAT AAAAAGTACA  
 50801 TCATGATTAT TGTATTCTT TATACTTGTG GAATTAATTG AAAATAGTTT  
 50851 TTATAGTTAA AGTCTTTCTT TTTATTGTT TACAGGCTGA AGAAAAGGCT  
 50901 CATTCAAGAGG TAAAAAAAAA TATGCAATAT TTTAATATTT TCTATTTAG  
 50951 TTTGCATTCA TGATGAAATT AGTCTTGTGA CCACTAGAGG GCTCTGTGAT  
 51001 ACAATAGCAG AACTCCACAG GACTGCTGAA GTAAGGCAGC TAATTGATAA  
 51051 ATGGTCTTG ATATTGCCTC TTAAAAATAA AATGAAAGGA AGTTTGTATA  
 51101 GCAAGCTGTC CTTTCACATT CTAGATTGAG TCTTAGCTCA ACACCTAATA  
 51151 AGTTTCTAT AATAGTAAGC ACTCATTAAG TCATTGATAA ATGAAGGTCT  
 51201 ATGGTCTTCC TATTTTATTA CAGTCTTTT CCCACTCCCT GTAAGACCAT  
 51251 CTACACAGGA TAATGGTGA AACTGGGCA CCAAGCCTCC ACAACACAGG  
 51301 ATACTAGCAT CTCAGACTAT CTGTTTGTG TCATTATCTT GTTGCCTCTA  
 51351 ACTGCCATT TATGTGTGGT GTGTCACCTA TTGTTCTAAT CACATATTTC  
 51401 ACAAAATACAT ATTTGGTTGC ACTCGTGAGC AAATCAAAC GCATTCAGGA

51451 AAGAATACTA TTTTAATTTC CCTTGGTAAA ACATTGTCC TGGTCAAAGA  
 51501 GAGCAGGAGG ACTTTAATTA TGACTTTATT CAAGGTGAGG TAATGGCTGT  
 51551 TTGATTGGTT TACACTGAGG CAATCAGACA ACAGAGAAAA AAAATGCCTT  
 51601 AACAAACAGCT TTTGCAAAAG TATTCCCTTC CTTTGAAGTC TTATTTTATT  
 51651 AGCCTTTAAA AATAAAATTT GTGCTATGTT TAAAAAATATT TGAAAATTAT  
 51701 TGATTAAACC AATTGTCTT TATAATCTCT GAACCAAAGA GTGGATATGA  
 51751 TTTTTAAAAA TCAAAGTGGT TTTATTACA TCACATGGAC ATGACAAAGC  
 51801 TTCTAACACT GATCATAGTA TAGCTACTGA AGCATCGAAA TGCTACATCT  
 51851 ATTTGCCTTA GTAGTAGTTA TTCAACTCCC CTTTTATCAT TGATGCTGTA  
 51901 TCATGAGTTA TGGTTTAAAA AAACAATTTC AATCACTTTA CAGTTTCCTG  
 51951 GATTATATTT TAAAGATACT GGAATCATGT AATAGAGACT ATTTAATTG  
 52001 AGAAATGCTC TTTGAGTTG GATTCACTTA TGAATAAAAT AGACGCTGTA  
 52051 TTTTCTGAAA TCATTCTAG TCATTATCTT ATAAATGTAA AGCAAATGTT  
 52101 ATTTTAGACT GGGGTGTATC TGTTCCGGAA AAAAAAAAAA ACAGGAACGA  
 52151 AGTAGAAATCA CATTGGTGA ATTATATAA GTGTCTACTG TTTCCAGCTT  
 52201 AGAGTTCTCT ACTTTGTTAG AGTGTGTTGAG TTGACCACCA TTTATTTCA  
 52251 ACAAAATCTA ATGCCCGGGG CAAAAACTAG ACAGTTAATA AACTATGTCA  
 52301 AGAATTCTCT TTCAAACTGA GACAGCATTG CAAAAGTTCA ACTACAACTA  
 52351 TAGATAAGAT TTGTTTTGA AGAAATGAGA AGCATCAAA GTAGAATGTT  
 52401 TAACATCCAA GTAAGTAAA TCCCTTGAGA CTAGATATAT ACTTATAGAA  
 52451 CCTAGTGTCA GATTGTTATA AATGTTCTAT CCTTATTAGT CACAACATGA  
 52501 GACTTGCAGA ACAAACTGCA GAAAGTGCCTT GAATTTAAAC TTTAAACATG  
 52551 ATATAATATA TCCTTACCTT TTTCTGTTTC AGTTTATTG GAGTGTGAAC  
 52601 TTAACTAAAA AGAAAGATAC CTTAGAATAT ACATTATATT GGTTTATCTA  
 52651 ATTAGTTGCA CCTATCATTG GTTTTTCCC CTGATTTTA AGATGTGGAT  
 52701 AAGCTATAAA GCATCTCTGA GCTAATAATA ACTCACTAAA TAAAGGTCTT  
 52751 GATAATACAG ATTTGGGAAG GCTTCTCTGC AGTCATTGAA ACTCCAGCCA  
 52801 ATAACAATTAA AATGTGAAC TGATTAAATG TTGAATTAAG CCCAAGTTT  
 52851 AGTGATTGCA GGATATTCCA TAGCCTTGA GAAGTTTCA AACTATGAGA  
 52901 AATTAAAATG TACAGAGGAA AAAAAACCT AAGATTTCT GAAAAAGAAC  
 52951 ATGGAGTATC TTTTACTAAA AAAGAACAAAG AAAAATATGT GTGTATATAC  
 53001 AGTTTTATA AAGAAAATAT TTTTCTACAG TTTTATTACC ACAGTTTTTC  
 53051 TAGAAGGAGA AGAATCAATA CAGAGGGTAA ACTGCTCTTG AGTCATTG  
 53101 CATTGAGGG ATGGCAAATG GAGCAAGTGA GCGTACTTTG ATTTGTAGAT  
 53151 TAGAGTTGA CACATAACAC TTTGCTTTG AATGACATTG GCTTGTACT

09/914549

35/124

53201 GTGGAGTCAG TGTTCATATC CTTTATTTC AGGAGTTGCT GCTGATAACAA  
53251 TGGGGTTAGA ATGAGCTAAA TACAGCATTG GCTTTCTTGG TTTGAATTCT  
53301 GGGTTTAAG TAAAAATCTA CTTGCCTATT CCATTGATTT TTTTAATTGC  
53351 ATTCAAGCAAA TCCATAAACT GCGGAGAGAG CTGGTTGCAT CACAAGAAAA  
53401 AGTTGCTACC CTCACATCTC AGCTTCAGC AAATGTAAGT CACTTCATT  
53451 TTAAAATATA TTACAACAAA TTTTATAGA GGAAAATGAA ATCATTAG  
53501 TAACAAACTT ACAAAATTTTC AGTGCCTGAT ACAGACTTAG ATTACCAACT  
53551 AGCAGGACTC ATAAAAAGTT AACATTTTT GCCTACTCAG TAATAAAATG  
53601 TAAATCCAAA CTGATGAGAG GCAGCAATAT GGTTAAAATG GCTTGTTGTT  
53651 TCTAATAAGA TTGGAAACAA TAGAACAGC CATATGGGTT ACTTCCTTTC  
53701 TTGTTGCTA TTTTATTAC TCCTCTTGCA TAAGATTCCC TGACAATGTA  
53751 AGAGGGGTTG TTAGTGTGTTG ACTTTGGAAG ATAAAATATT CCTGTGCCA  
53801 GCCTCCTTCA TCTCAATGTA TTGAACAATT TGTTAAGCAT CCAGTTAATT  
53851 CTAAAATATG AAATTAGGTC TAAATAGGGA TAGCTTAGCT GCACTGTGGA  
53901 TGAGATATGG TTTGCTCAAA AAACCTTGGC AGCCTCTCA TAGCAATT  
53951 AAAGGGTACA CTTTTACTGG CACCAAGAGCA GCCCAGGATG GCAGAAATGA  
54001 TGACAATGAA GACCGTCAAT TAAATTAACA TTTACTGAAT ATCTTCCACT  
54051 GTGTCAGGGA GCACTCAGAG TAGATGCAGA ATGATAAAGG AGAAATGTGG  
54101 CACTGTTCCC AGTCCTGAGG AGCAATGGTG TTAAGAACAG CAGTGAGGG  
54151 TAAGGAAATG CCTGCTATTT TGCCATATGT CTTACCTCTC TCACTCAACA  
54201 GTCCTTGCT CAGTTCTGCT GCATAGCTTT GGGCCTGCTC TGTGCCTCCC  
54251 CACCCCTCCC ACTGCTCCTC TACTGAGTT TTCTATCTCC TAGACAAAGC  
54301 ATGATATGTC AAGAGTGAGC AGGTGCAGAC CCACAGTGT AAGACTTGAA  
54351 AAGAGCCATT TTTAAATT TTTAAAGCTA TCATTGTGCA ATATAAATTC  
54401 TAAGTATGTG TATCATTCA TTCACAATGT ATTCAATTAA GCACTGTATT  
54451 TGAATTGATT TTATTTCTG AAATTTGGGA GAATTAATT TGGATTTATT  
54501 CTATTTATT TTAATAGATG GTGTTAGGAG ATTCTGAAA ATAATAGCAG  
54551 TTTTAGATA ATTGTTAAG CAATATGAGA AAATAAGGT ATTATTTAAC  
54601 CTTGTTGTGT TTTAAAGAG ATAGTCCAGA GGCAACCGTA AATTTTATAA  
54651 TATAGGCTAC ATGTATAGAA GTATGAAATA TTGTTGTCTA GGTCCTGAA  
54701 TTTGTACCCA GAGGAAGTAG AATAATGTAA ATGTCAGAAC CTCCTGGGTT  
54751 GTGTTATCT GCAATAAGAA AGGCTCAATG GCAAACCTTA TTTATTAGAT  
54801 TGTCAAGGATA CTTGCAGATG TCTTGAATGA TTACTCAGGG TTTCATT  
54851 TTTTAATGT CCCTGGTTG AGCTCATCAT ATAATTCAAGA TATTGGAATA  
54901 ATAAATGGCT GCTAGACATA GTGGAAGATG GGCTGATACT TTCCATT  
54951 AATGTAATGA TGCTTATTGT CTTCAAAAGA AAAAACTAAA ATGGTATTTC

Fig. 2 (cont'd 30)

36 / 124

55001 ACATTTTTT GTTTTGTTT TTGTTTTTT TTCTCTGAGA ATCTCATTCT  
 55051 TACTCATGAT TATTGGTTTC TTGTGTACCA TTTCAACATT TTTCTATTAT  
 55101 ATGCTAAATGT GTATATATAC TTAATACACA CGTGCAAAAG CTTCCACACA  
 55151 CACACACACA CACACACACA CACACACACA CACACATACA CACACATACG  
 55201 GAACCAAATT CTAACATAGG GGAATAATCT TCGGAGTGAA CTCTGTGCTG  
 55251 CTGTTTGAAA ATGGAGATAT AATTTTAGAA AGGTTCTGC AGTTGGCTAC  
 55301 CCACCTCGTC TGCTCTAATT ATGCTTGTCA CACTATTTTC ACTGATGTGT  
 55351 TTTCATGACT TTAGGGCATG AATTCTCAGC TGGGTGTTAA TATGACCAAC  
 55401 AAAGGGTGAA AACAGGTTCT TGCAATTTTT TAAGTACTCT TTTTATGTGA  
 55451 AAAGCACAGA TATGCAGATA ATACATAACT GAACATCCAG CATATCTGTG  
 55501 GCTTTAAAAT ATCACGAAGA AGAGCACAAT TAGGGAAAAG AAAACATCTA  
 55551 TAGTGTTCCTC CTAGGGAAC AATCATTAA AAAAAAATAA AAATAAGGAA  
 55601 CACAGACTAG AAGCAGCAGT GCCAAATAGA TAATTCTGC TAGTCTTTGT  
 55651 GTTAATTAA AAAGTGCTAG TCTTGGAGAC AAACGCCAA ATTGCTCTAG  
 55701 GTTCCACTCA GCTGTATGTG TTATCATTAG TATTAACCTT TGCACGCTGA  
 55751 TGGGAGACTG ATATATATCC TGTTTTATGT TCCTTTAAAC AATTTATAAT  
 55801 GTAATTAGA AACCTCTCA AATCACATTA GATCCACACA AAAACCTGTA  
 55851 CATAGCAGCT TTATTTTTA ATAGCAAAG AAAGGAAACA ACCAAAAATA  
 55901 TCCCTTAATA GGCCAGTTAA TAAACAAATT CTGATACATC TATATCATGG  
 55951 ACTACTACTC AGCAATATAA AGAAAATGACT ATTGATACGT GCATCAACTT  
 56001 GGGTGGATCC CAGGGTATT ATGCTGAGTG AAAAAAGACA GTTATAGAAG  
 56051 GTCAAATTAA GTATAATTCC ATTTATATAA CATTCCAGAA ATGGCAAAT  
 56101 TAAAGAAACA GAGAACAGAT TAGTGATTGC TAAGGGCTAA GGATGAAGGA  
 56151 GAGAGAGAGG TAGTGTGACT ATAGGAAGAG GGAGATCTTT AGTTTTGTAT  
 56201 TTTGAATGAG ATGCCATCA CATGAATCCA CATATGTCAA TCTATTAATG  
 56251 TAAATCAATA TTGTATTCTC GGCTTTGATA TATAATATAA TTTTATAAGA  
 56301 TATATAATCA TTGGGGAAA CTGGATGAAG GATACAAGGG ACCTCCCTGT  
 56351 ACTATCTTG CAACTCTTG TGTATATAAT TATAAAATAT ATAATGTATT  
 56401 AAAATGTATA AAATAATATT TTAAGTATCA GATACTGATC TTTACTCAGT  
 56451 ATATGAAGTG TTCTATCATA ACGTAACATG CTTTTCCCTTT ATTTGTGGTA  
 56501 TTTTAGTTTC AAACTAAAAT ATAAATCACC TAAAGATCTA CGACAGTTCT  
 56551 TTTGAAAAAA AATCTTGCTT TTAATTTCCC AGGAGTTCA ACCTTAATCC  
 56601 TCTCTTTAGT GTTTCTTTAT TTGGTAGTGA TAGGGACTAT CAAAGCTTCT  
 56651 TACCATCAA TACATTTACT GACTAAAAAT AGAAAAATAA TTACACATTGT  
 56701 AAAAATGTAC AAATTGAATG ACAGTCAAA GGTACAGGTA ATGAAGATAT

37/124

56751 GCATTAACAT CTACTTTAA AAAAAAGTTT ATTAAAATTC TCTTTTAGAC  
 56801 TAATGCAGTA TCTGGGAATT TATATAAATA GATATGTATA TAAATGACTA  
 56851 TTAAACAATT TTAATGTCAG TTATATTTA AACATTTAA TAATATTGTT  
 56901 ATAACATATGG GGGTAAAATT TTGTATATAT CTGAACATTT TTGTTCTTAA  
 56951 GGAAATAATC ATTTTACAT ATCCAGGAAT TTGAATTACT CTCAAGTCAC  
 57001 CTATTAATTA CAAGTCATT TGAACTCATT CATTCTTT GTGTTGCTT  
 57051 TATAATGTCA TTTTAGATTT CATGCATCAT AATCAGCCAT CAAATAATTT  
 57101 AGTTAATACT TGATTTTCC TCAGTTGTA GAAGTGCTGT GTTTAAATTT  
 57151 CATTCAAGAAT GTTTCATTTC ATCTGAATTA ATATCTGTTA ATGTATGTAA  
 57201 TATACACATA TTTTAACAT GCATGTACTT AAATTGATTA TAGGGACTTG  
 57251 GTAAAATTAC TTATTTATAG GATATTTAA ATATAATCAA GGATTTTTA  
 57301 AATCTACAGT TCCCATTGA AAGTAAAAGT AAGTCTTGT TTACTAGTTT  
 57351 GTTCACAGTA CAAGTAAACT TTCTACCTT TGGTTAAATG TGAGTGCAGC  
 57401 CCCCCACAGTG AGAAATTGTT ATATTAGAAC TCTAATAGCT ATAATTTATA  
 57451 GGGATGAATT TCAATGAGTT TGGTTCTAAG AAATAATCTG TTGGTTTAA  
 57501 CAACATTTT AAGTATCAGA TATTCTACATT TACTCACTAT GTGACATGTA  
 57551 CTCTCATAGC TTACGTGCTT TTCCCTTATT TGGGGTGT TTTATATATT  
 57601 AATTGGTATA TCGCATATTT AAACCTGGCA TAATTACATT TATATGGACT  
 57651 CTAAACAATA ACTTGTATTT TAATTTTAA ATTTGAAATG CATCTATGTC  
 57701 TCTGTTAAA TGCATTTCTT TCCCTTGCC CAAATGGGT ATGGTAAGTC  
 57751 AAGAGAGTCT CTAGTTAGCT CACCTCTCAT TTGACTGGCA GAGTAAAGCC  
 57801 CTTGTTCACT AGAATGTGTG TTAAGCCTTC CCTCCCTTT GTAAAGTTGT  
 57851 TCTGAACAGA GCTGCATAAA ACCACAGGTA AAGTGTAAAG CTGATTCTAC  
 57901 TAGCATGTCC TTAGAAAGGA GAGCGGTTAT ATTGGCAGGT CCTATTGCCT  
 57951 GGCCTTCTG ATCAATAACT CACCAACAAA CAGAAAACAG AAGCCGCACA  
 58001 AGGAAAGGCA GAACTAAATA AATGGTAATA GCAAACAATA AGCCAGATA  
 58051 CCTCTGGCCT CTCGCCACCA CCTTAAGGCA GCTGGTCAG GTGGGATGCT  
 58101 TTGTTGTC TTTAACGTA TTTCTTAC AAATCTCAGC CATTACATAA  
 58151 TTTGGAAATG GACACAAGGC TAGTTATTAC TAACATTTT AAAGACATTA  
 58201 CTGAATGAAT GTGTAAGAAA ACAAAAGGTC CTTTTGCTT TTCAGCAGAT  
 58251 AAGTCTTTA ACCAAAATC TCTTGGGTAT TTTGAGATTG TGTTCTACTT  
 58301 CTTTGCTTAT TTAATATTTT CATAAAATTT GCTAGTTACT CTTGCTTTT  
 58351 TGCATCTCTT CTAAGAGAAA ACAATTGGTG CATATTATTA ATGAGAAACA  
 58401 CTTCAGTGTGTT TGGACAATTT TTTGTAGTGG AAAAGAAATG TGAAACTTTA  
 58451 TGTTGCAGAA TCATTCTTGG TTCAACTAAC TACTAATTTT AAAACATATAA  
 58501 GTCTTAAATA TATATAAAGT TTATATGGGT AAATATATAT TACATATAAT

58551 ATATGTTTA TATTATACA TAATATACTA TATATTATA CATGATATAAC  
 58601 TAAATATTTT CCCATATAAA TAATAAAATG CTCTAGGCAT ATATGTGTGT  
 58651 GTGTGTATAT ATGTATATAT ATATATACCT TCATAACATA CATATATAAA  
 58701 ATACTATATT ATATATACTC TAGGTATACA TATATGCCTA TATATGCACC  
 58751 TATATATTAA TATATTACTA TATAATATAT AGTATATATT ACTATATATA  
 58801 CTACTATATA TTACTATATA ATATATAGTA TATATATAGT ATATATTATA  
 58851 TAGTAATATA TTACTATATA ATATATAAAAT ATATGTGTGT ATATATATAT  
 58901 ATGCCTAGAG TGTTTTAAT TTGTCAGTGG GCTGTCCTG TAATCTATAT  
 58951 GAAGAAATAA AATGTAGACG TTATGTATAA TGATATTCA TCTTGTGTG  
 59001 TGGCATCATA GTAATCTCT TTACATATCT ATTCAAGATTA CTTTGCACC  
 59051 AGCCTAACATC ATTGTATGAT TCCAAAACCA AAGAGAGTAT GGATTGAAAT  
 59101 GATATCCCT TTACTAACATC TCAGTCTTGT CTATTTATT ACCTTTATAG  
 59151 ACTTCACCTA ACACAAGTCA GGGGATATTT ATCATCATAT TAATACAATT  
 59201 TTACTCTGAC CTTAAAATTA TGCAACTGCT AAAGGAAAAA TCAGAACCAA  
 59251 ATAAACTGTC ATTAACAACC CCCCTGAAAA TCCATATTTT TTAAAAGTCA  
 59301 TTTTATCAAG TCTCTCAGAC AAGATGTGAT ACCCTATAAG TTTAATCAGT  
 59351 TTTACTTTCC ATTTCTCTT CATTAAAGGTG ATAAAGATTA TCATTAGTAG  
 59401 AAAAATTTTC CCTTATTTGC CTCCTTTCC ATTTACCTA TTGAGTGAGA  
 59451 AATTTAGCCT CTCATAACTT CTAAAGTAGC AATGTTAACAT TGATAAACTA  
 59501 AACCAAGGTG AGATAAAATTT AAGACAATAT TTTTTTCTT CAACTTTAA  
 59551 GTTCTGGCGT ACATGGGCAG GATATGCAGG TTTGTTACAT GGGTCAACAT  
 59601 ATGCCATAGT GATTGCTGC ACAGATCAAC TCATGCCTA GATATTAAGC  
 59651 CCACCATCCA TTAGCTATTC TTCCTGATTC TCTCCCTCCC CTAACCTCCC  
 59701 CTGACAGGCC CTAGTGTGTG TTGTTCCCCA CCATGTGCC ACGTGTTCTC  
 59751 ATCGTTCTAC TCCCACTTAT AAGTGAGAAG AAGTGGTGTGTT TGGTTTCTC  
 59801 TTCCTGTGTT AGTTGCTGA GGATAATGGC TTCCAGCTCC ATCCATGTCC  
 59851 CTCCAAAGGA CATGACCTCA TTCCCTTTTA TAGCTGCATA GTATTCCATG  
 59901 GTGTATATGT ACCACATTTT CTTTATCCAG TTTATCATTG GCATTGGGT  
 59951 TGATTCATG TCTTGCTAT TGTGACTAGT GCTGCAGTGA ACATAATGCA  
 60001 TGCAGGTATC TTTATAATAG AATTATTTAT ATTCCTTGG GTATATACCC  
 60051 AGTAATGGGA TTACTGGTC AATTCTGCT TCCAGATCTT TGAGGAATCA  
 60101 TCACACTGTC TTCCACATTG GTTGAACCAA TTTACTCTCC CACCAACAGT  
 60151 GTAAAAGCAT TCCTTTCT CTGAAACCTC TGCAGCACCT GTTATTTCTT  
 60201 GACTTTAATA ATCACCATTC TGACTGCTGT GAGATGGTAT CTCATTGTGG  
 60251 TTTTGATGTT ACCCTTTTT TTATATGTTT GTTGGCTGCA TGACTGTCCTT

Fig. 2 (cont'd 33)

60301 CTTGTAAGTG TCTATTATA TCCTGTCTAT TCATGTCTTT GCCCACTTTT  
 60351 TAATGGGAA GTTTGTTTT TACTTGCGCA TTTGTTGAAG TTCCCTTGAG  
 60401 ACTCTAGATA TTAGACCTTT GTCAAATGGA TAGATTCCAC AAATGTTCTC  
 60451 CCATTCTGCA GATTGTCGT TCACTCTGAT GATAGTTCT TTTGCTATGC  
 60501 TGAAGGTCTT TAATTAGATC CTATTGTCA ACTTTGCTT TTGTTGCAAT  
 60551 TGCTTTGGA GTTTTGTCA TAAAATCTT GCCCTTACCT ATGTCTTGAA  
 60601 TAATATTGCC CAGATTTGT TCTAGGGTTT TTATAGTTTT TGGATTTAC  
 60651 TTGTAAGTCT TTAATCCATC TTGGGTTAAT TTTTGTATAA GGTATAAGGA  
 60701 AGTGGTCCAG TTTAATTCTT CTGTATATGG CTAGTCAGTT CTACCAGCAC  
 60751 CATTATTAA TTGTTTTTC AGTTCCCCA TTGCTGTCTT TTGTCAGGTT  
 60801 TGTCGAAGAT CAGATGGTTG TAGGTGTTT TCACTAACAT AATCATAACA  
 60851 TACATTCAT TGAAAACAAC ACGACTCAAA ATGTTCTTA GTAACCAGTT  
 60901 ATAAGTTTT TTGTGCATAA TTACAAACTG CCATTCTAAT CATAAACATT  
 60951 TTGTTGGTAC TTATAGCTAG AAAATGTGAG TAATATAGTT TATACAGCAT  
 61001 ACTCTTACA ATCCCATT CTTTGTCAAA CTTTAATTCA TATTAAATTG  
 61051 ATAAAGTATA CACAAAGGGT AAAGGAGAGT AATTTCTTC AAGTTTCACA  
 61101 TTTAAGGATT CATAGTAGAA TGATTAAACC TTACATTTCT CCACTATAAG  
 61151 GAGAATTAAA ATGAAATAT TGAGTAAAT CTTACATTTC ATTTAGTAAG  
 61201 TGCTAATAAA GGGTTCTGC CATAATTTTC CTTATTTAA AAGAAAACAC  
 61251 ACAATTTCAG TTTTAGGTTT TAGTAACCAA TTTTATGGC ATAGGGAA  
 61301 TATTTCTAAC AGGTTAAACT GAAAGTGACCA TCATGGCAT ATATATATAT  
 61351 TTTAAATTCA CATATATGAA TACTATACAG TAAAAACTAA CTTATGCTAC  
 61401 ATACCACATG GATGAATCTC AAAACCCATG TAAAGCAAA GAAAACCACA  
 61451 AAAGAATCAT GCCATTTGAT TACACTGGG TGGTTTTAA AACAGGCATA  
 61501 TCTAACATA GTGCTTTAAA GTGTAAGCTT GGGTAGGAAA AACTATAAAG  
 61551 AAAAGCAAGA AAATAATTAC CACAGAAGTT ATGTAGAGGT TATCTTGGG  
 61601 GAAGGAAGAG GGAATAATAA GAGAGGGACA AAGAAGAGCT TCTTGGTTCT  
 61651 TGAAATGTCC TATTCCTGCA CTTGGCTGGT GAATGCATGA ATGTTCACTA  
 61701 TGTGATAAGT CAGGGGCTG TTTTCATTTC GTTCACCTTT ATATATGTGT  
 61751 GGATTTTCC ACAGTTGAAA AGGTAAAGTT CAGGTGTGGT GGCTCACACC  
 61801 TATAATCCA GCCAACACTT TGCAGGGCCA AGGTGGGAAG AATTACTTGA  
 61851 GGCTAGGAGT TGGAGAGTAA CCCAGGCAAC AGGGTGAGGC ACTGTCTCTA  
 61901 CAGAAAATGA AAAAAAAA AAAAAAGTAG CTGGCATGT TGGTACATGC  
 61951 CTATAGTTCT TGCTACTTGG GAGGCTGAGG CAAGAGGATC ACTTTAGCCC  
 62001 AGGAGTTAA GCCTGCAGTG AACTAGGGTT GTGGCACTGC ACTCCAGCCT  
 62051 GGGTGGCAGC AAGACACTGA GTAAAAGAAT AAAATAAATA ATTAAAAGTT

40 / 124

62101 AAAATATAGG AAAAAATGAG CATAGCCTTA TGCTAATTT TCAGTTACTA  
 62151 GGTCTGATAT CATCACATTC CTTGCTTGTC ATTGAAAATT TTTTAAACTA  
 62201 TGATACTTTT TTTTAGTGGT ATTTATCCAA TTAAATCTGC TAACAAATT  
 62251 GGTGTATAAA TCTCAAGGGT AAGGGTATGT GGAGAGTGGG TGTGTTGTG  
 62301 TGAGAGAGAG AGAGAGAAGA GGGGGAGGAG AAAAAGAAGG AAGAGGGAAG  
 62351 GAATGGAAAA AGATAATAAA GAGTTGTTCT GATAGATTAA TCTTTAGTAG  
 62401 ATGTATTCCC TACAAATTGT TTTTCTCCAT ATTGCAGTGT CAGGTAAAGA  
 62451 AAGGCATCCC AGGATGAATT CAGAGCTAGG AACATGCACC TTTGTATCAT  
 62501 AATGCTAATG GAAGGAACAT GTACATTCTA ACTGTTACCA ATAATGGAAT  
 62551 ATATTTCCGT TATTAAGTAA TAAGCTTTAA TTCTTGTAT TTTTGTGATC  
 62601 CATTGATAG TAGGTGCCTC AGCATTCCA CTCTGCTATA AGTACATGGA  
 62651 GATATATTAAATTTT ATTTAAGTCA TCTTATTCCAT GTCTTCAAA AAGAAATTCA  
 62701 TTTTTGGCCA AGGATTTCCA AATTTGCCA CATATATAGG TATAGTTTAT  
 62751 TATAGACTTC GTTTGCAAAA TATTAATCC TTATATCCTT TTAGGGACAC  
 62801 AATAAAATT TATAAGTTTG AGATAATGTA CTTGCAGTTC TACCTCAGGC  
 62851 CGTGGTGAGA GATTGAAGTG CCTCTTCATT TTAACATTTT GGGTTCAAGT  
 62901 TGTTGCATAA GGGCATGCAA ATGGAAACTG GCCTATTTT GAGCTTTAAT  
 62951 AAAATCGTCA AATACTTCTT AATCTTAAGA GTTATAGTTA TGTACTACAA  
 63001 TATGTATAAT TCTCTAATAT TTAAAACAAA ACCTGAAAGC CACAAAGCT  
 63051 TACTGTAAA TAAAATGTGA TGGAATATTA TTTCTAATG GCTTACCTGT  
 63101 ATTTCTTCA TTGAAGGGAA TATGAAGTAG AAAAGCCCTT TTATTGAAAA  
 63151 GAGTTTGAGA AGTAAAGATA ACTCTTTCA ATTCAATTCT TTGTAAGTAG  
 63201 AAAAAGAGTA AAGATAATGT TTAGCTGTCA GCAGATGTCT GACACTTGAT  
 63251 GGAGCGTATC ATTACAATAG AGCAGCTAAC AATATCTGCA AAGGTCATCA  
 63301 TGAAAGTATA AAAATGAGGA ATATTTGTCC ATTGACCATT TCAGTGACCT  
 63351 CTTTTGGC TTTAAGTCTA AAAATCTTGG CAGATCAGAA CTTTATATTC  
 63401 GGCATTTGAGA GTGTCAAATC TCTACATGAT GTGCAAGTCA GAAGGAGTTA  
 63451 TTACTTGCAA AATACCATCT TCTTTCAGAA GTTAAACTCA CATTAAATGC  
 63501 CAGGAGACTG AAACACTGAT TTTAAGAAGA CAAAGTTAG AAAAGATGAA  
 63551 TGAAAATGTG TGTTAAAGAA GAGTCACCAAG TCAGAGCTAA CTATGATAGT  
 63601 CATAGTATTT AAAGAGTTGG AACACATGAA ATTAAGCATT TTGTAAGATG  
 63651 AAGGCTTTTC ATCCATCCAC ATAAGATTCT GACATTTAAA CTATGTTCT  
 63701 TCCATTCTGT TCACAGGCTC ACCTTGTAGC AGCTTTGAA AAGAGCTTAG  
 63751 GGAATATGAC TGGCCGATTG CAAAGTCTAA CTATGACAGC GGAACAAAAG  
 63801 GTATGTTCAAG AAATTGCCAC TGGAGACTGA AAGAAGACAG CAAATTGCAT

63851 AGGATTCTTA AATAATACCT GAAGCTCCTT AAAAATAATA TTCCAGGCTG  
 63901 AGTGCAGAGG CTCATGCCTG TAATCTCAC ACCTTGGGAG ACCAAGGTGG  
 63951 GTGGATCACT TAAGGTCAAGG AGTTCGAGAC CAGCCTGGCC AACGTGGTAA  
 64001 AATCCCACATCT CTACTAAAAA CACACACAAA AAATTAGCTG GGCATGGTGG  
 64051 CGGGTACCTG TAATCCCAGC TACGCAGGAG GCTGAGGCAG GAGAATCACT  
 64101 TGAACCCAGG AGGCAGAGGA CGCAGTGAGC CAAGATCACA CCACGTGACT  
 64151 CCAGCCTGGG AGACAGAAC AAAAAAAGAG TAATAATAAT AAAATAATAT  
 64201 TCAATTCTAT ACTAAATTAA ACAATGATA ATACCTTCT TTTCAGATTT  
 64251 TAATTTAAAG ATTTTATCAG TTTACTCCAT ATTGGAACAC ACAAAAGGCAA  
 64301 ACAAAATCCT TGCTGGGCAG TCTATTAAATT TACTTCTGGA TGGAACTAGT  
 64351 AAAAGAATAC TGAATGTTAA GAAAGAGAAA CAGTCACATA AGAGAATATT  
 64401 CTGGGGGCAA ACTGTTATGCA AGTTGACAAG AATCACACTT TGATAAGAAC  
 64451 TTTCACAAAT ACATGGTCAC TAAATCCAGC TATAGGCAT GGCTGTAGGC  
 64501 TAAGACACAC AGGAAGGATG CCTGGGACTC TGCCAAGTAA GGGACTTCAG  
 64551 GTTACAGCAG CTATGAAACA AAGGCCAATC CTGTGTAATT TTGAAATAAC  
 64601 AAGAACTAGT TGCCATCTAG GGATATCACC TTTGAAGAAA AGTCATTG  
 64651 TATATCAAAA TACTTAAAT GAACCTAAAG GATTTATGG TATGAAAGAA  
 64701 GGTATACCAA AAAGAAAGGA ACGGAGAATT TAGTTCACGA AGACAAATGT  
 64751 ATTAAAAAGG TCCATACTGC ATAGAAAGCC TGGTCACCTT TCCTGTGATG  
 64801 ACCAGTTAGC TTACTTCTCT GCTGTTAGTC CAGTGGCCTT AACTCCCTTG  
 64851 GATAGGTATC AGAGATAGGT GAAACCTATA GAATTCTATG GAGTGTGTGT  
 64901 GTGTGTGTGT GTGCGTGCCT GTGTGTGTGT GTGTGTGTAT GAAAAGTGT  
 64951 AATGTGCATA AATGATCAGG TGTCCAGAGC TTTCATCTAA TTCTCAAAGA  
 65001 GACCCATTAT ATCAGAAGTT TTGGGTATTT TCAAGAATGC GTTCCTCTAT  
 65051 CTATCCATAG GAATGGCTTC AGTTTGTCT TTAGATTCTG TAAGTTATGT  
 65101 GATTAGCTTT ACAAAAGTAG TATGTATTAC CAAATTGTGT CACTTACAA  
 65151 AAGTTTATTT TTAAAACAGA ATGAATAGTT CAATGAAATC AAAAGAGTAA  
 65201 ATCGAATATT CTTATAATTG CCAACTATTA TTAGCACATT GTATTCTCTC  
 65251 TCATATTCTC CGTATACCCCT GCCCGTGAGA GAGAATATTA TCCATTCCCTG  
 65301 GAAAATCTGT TCTAGCACAG CTAACAAACT CCTTTGAAA CATAAATTTT  
 65351 CCTTTCTTTC CTCCCTCCCT CCCTCCTTCC CTCCCTTCCT TCCTTTTCC  
 65401 TTTTCTTTCCTT TCCCTTCCTG CCTCTTTCT ATCCTTCCTT TCTCCTCCCT  
 65451 TACACCCCTT CTTCCCTCTT TTCCCCCTCT GTCTCCCTCT CTTCTTTTT  
 65501 TGCTGCAGCT TGTCACCTCA CTATGTAATA TAAGAACCCCA GCAAATAGAA  
 65551 TTAGAAGGCT TTTTAGAGCA GCTGACGGGA AAGAATAAAA ACACGGCCC  
 65601 CCAGTATTCT TGAATGAGAA TTCTGGCTAT GTCTGTTAAA AGCTGGGTAA

42/124

65651 TCTTGAGCAA GTTTATCTAA CCTTTCTTGA ACCTCAAATT CACCTTCTTA  
 65701 AAAGTGGGGA TGATAATGAC TACCTTGTAG GATCACCAG AGGAGTAAAT  
 65751 CAGATACTGT TATCATGTCA CATGCTAGGG GCTACCAAAA AATATTACCT  
 65801 TCCTTACAT TTCTCTTTT CCCTGAAAA TTATAAGATA ACACCAAATT  
 65851 CCTCACTGGG CATATACCAA GCATATTGTT GGAAATGAGT CTTAGAATT  
 65901 AAGTCTCAAT ATCTTTAATA AGTCAAAATT AATAGAATT TTGTCCTCCA  
 65951 CCCAATATTT TCTTGAACTC TGTTATATCT GTAAGTGAAT TTTCTCATAG  
 66001 AACACATACAG AGAATTTCT CATATACATA TAGAAAAAAA TGTAGAGGTA  
 66051 TGTAAATGTA TAATGCCTAT GATTAATGCC TGAATATTAA AAAATAATT  
 66101 CTATAACATA AGAGATTTA TAATGTGTCT ACATAATCCT TAAAATAACA  
 66151 TTGCCAAAAT TATAAAATT TCTCAGAAGA TATCAGAATG TCTCATATTG  
 66201 TCCTTACAC TTTTTTAACT GAAAATAAAA TCACTCTTT TTGAATTGCA  
 66251 AACTGTATAC ACACAACAAT CATGGTTAAC TAGTTTATTAA ATTTGAGATT  
 66301 ATAACCTGCC TATTCTCAAA GTGATATTAA AAAGCCTATA AAATTATTG  
 66351 CAATGTGAAA TGGTATAATT CAAAGACAGA ATCTAATTAA AACCAAGTAGA  
 66401 ATAATGTATA TAACAATATA CCTCAGCCTA GATAATTACT ACTGCAAGGC  
 66451 ACTGAAATGA ATTGAATTTC AAGGAAGCTA TGGTACAAAG GGAGATTGTT  
 66501 AGGTGTGTTT TATTCTCATT TTCTGACCAG GAGAGCATAA TTTAGACTGA  
 66551 GGAGAAAAT CTTTGGCACT AAATTCAAGG ACGAATTAT TGCCAAGGTT  
 66601 TTTAAATGG GGTCACTGGAA TAACAAAAGA CAAAATCACT GTCAAATAG  
 66651 ACATTTCTCT AAAAGCTAAG GGCATAACAT TTAATCATAT TTCACTAAAG  
 66701 GCATTTCTTC AGGGAGCTGA GATAAAAAGGG TATATTGCTC TCTGGTGATT  
 66751 CAACAATCCT GAGAAAAGGC TTGTGAAGTA TAGAGCAGAG ATTCTTAAAC  
 66801 TCCCTTCCCC AAGTTATAAG TTTCATTTGT CTATATAGTC ATTCACTCAAG  
 66851 TTTATATGTA ATTTGTGCTC TTCTAATGAC AAAACAGTAC AGACAATATA  
 66901 GATATAGAAT GATAGATATA GGTCTATATC TATAGACATA CCTATCTACT  
 66951 AGAACTCTAA AAGCATATTAA TACATGTATG TAATATTCTT CATGGAGTTT  
 67001 ATATTTCTCA TATATATCTC ATATATATGT ATCTCTTAT CATGGAGTTT  
 67051 ATATTTAGG AGGTACACAGA TGATAATAAA AATATAATTAA AACAGGCCA  
 67101 GGTGTGGTGA CTCACACGTG TAATCCTAGC ACTTTGAAAG GCCAAGGCAG  
 67151 GTGGACTCCC TGAGATCAGG AGTTCAAGAC CAGCCTGGCC AACATAGTGA  
 67201 AACCCCCATCT CTACTAGAAA CAAAAATTAG CCAGGCCCTGG TGGTGGGCAC  
 67251 CTGTAGTCCC AGCTATTCAG GAGGTTGAGG CAGGAGAAC ACTTGAAACCT  
 67301 GGGAGGTGGA GGTTGCAGTA AGCCGAGGTC ATGCCACTGC ACTCCAGCCT  
 67351 GGGCACACAGA GCAAGACTCT GTCTAAAAA AAAATATATA TATATAATAT

Fig. 2 (cont'd 37)

67401 ATATAATATA TATATAAATA TATATATTAT ATAATATATA TATAAATTAC  
 67451 ATATTTATAA ATATGTAATT TATATATATA ATATATAATT AAAACATATA  
 67501 GGATTCAGG TGATGATAAG CACTACTGAA AAAAGTAAAG CTGAGAATGA  
 67551 GGATACTGAG AAGCTGGTT GGAAGCTAAA ACACAAAGTA ACAAAGGCCA  
 67601 AGGTGGTTAC ATGTTCTTGA TTACATACTT TAAAAATGGA TAAACTAAAT  
 67651 TAAGACTCAG ATTCTAGTCT TTGGGCTTCA CAGTGTGATT TTCAGCAATC  
 67701 ACATGGCATT AATAGCCTGA AACTACATCA AAATTGTCAT TTGATTATA  
 67751 GACCAAAATA ACTCCCTTGA ATAGAGAGGG ATTCACTCCT AACACTTTTC  
 67801 CTATTTCCAG ATGCCAAATA ACACGGAATC TCTTGCCAAA TTTGTGTGGC  
 67851 AGAACACTGG TTTTATATAC TTATAGCCTG GTAAGAAAGA AAAGACATGT  
 67901 ATGAATAACT TAGAAGGCAG AAAATTATCA TGCTATTAGA CTCAGTACAA  
 67951 TGTCATGTGC ATTCTCAAAG GAAACATCTG CAGAGGCAGG AGAATTGCTT  
 68001 GAACCCTGGA GGTGAAGGTT GCACTGAGCT GAGATCATGC CACTGCACTC  
 68051 CAGCCTGGT GACAGAGAGA GACTGCATCT CAAAAAAATA AAAATTACAA  
 68101 AAATAAAAAA TAAAAAAATAG TGATCAATCT GGCAGCATT TCTGAAAGTT  
 68151 AAGCAGTATT CCCAATAGCT GCTAAAAGAA GACATGTTAT ATAATACTAA  
 68201 GTCTGTAAGT AGGTAAAAT TAAGAGAATT GTTAATGTGC TTGCTGGGG  
 68251 GTGAAATTAT CTCTAGGCAT TACCTATAC CTAACCTAGG ACTCAGTAGA  
 68301 CTATGATATT GGCGTAGTTT GACCAAGAAT TTTATCCTGA TTTCAGATCG  
 68351 TTTTCTCTTC ACCAGCACTT CTTCACCAAGG ATTATATGAA AAAATTAA  
 68401 CCTGATGCC TGAGGCATCC ATTATATGTG CTGAAATAAC TTCTTTCTC  
 68451 ACCATCTAGA ATGGTACTAG CTATGTACCA CTCTGTGAG AATCAAGGAA  
 68501 ATTGCTACTC AAATCATTGT GCAGCTTAAT TTTCTCACAG AAGGCCAGTT  
 68551 GAGAAAGGCT CAACTCTAG GAATCCAGCA AACTATATTT TTTATAAGTA  
 68601 ACATTTTAC AGAACTACTT CTAAATCCTT GTGTTCAAAT TTACTAAAGC  
 68651 TATATTCAACA GCTAAATATT TCAGAATTAA AAATTAAAAA GACTTTCAAA  
 68701 TTAGTCCCT GTAGCTGTCA TGCCAAGGCA ATTAGAACAT ATGTTAAGGT  
 68751 ATGAGGGTT TTTCTGTGTA GAAGGTAGA GCAGGGCAGA GAAGTAGCCC  
 68801 CTTGTATGAG TGATGAAGCT CAGATATTGA CTCCTATGCT AACCATAAAG  
 68851 CCTAGTAGTT TGCTCATTTG TTACCTCTCT GAAACATTTT TTTGGGTGAC  
 68901 TACAAAACAG GAATTGAAAC CTTCAAAATA AGGGAATTG AAACCAAATC  
 68951 TTTGAAAATA GATAATGCTG CAACTAAAAA TTTAGTTGAA TAAGATTTT  
 69001 ACATTAACTC TCCCTAATT ACAGTTATGAT ATTTGCCATC TAGAAGTGTT  
 69051 TTTAAAAAAAT ATATTGCTGG AGTCAGATGA TGCATCCATT AATCTTTGGG  
 69101 GCATAGAATA ATGTGAATCT AAAATTTCA AATTATTTAC ACTACTGGTA  
 69151 TTTGGTCAAT GTAATTATT TGAAACTAGA TGCAATAGGG ATGGCCAGGT

09/914549

44/124

69201 TATTCAGTA GAACAACTAG CAAGACTTCA GATGCATGGT GGAGTGGGA  
69251 AAGGAGGACC TGTTAAGGA AACTAGAGCT GGGAAAGTGTG AGATTAACCTT  
69301 AGTGCCAATG TGAGGACCTA AAAAGCAGAT GTGGTGGAAA ATTTAACAG  
69351 GCTTGCCTAG AAGGTCAAGT TAGTTGATGA CACTTGATGA GATTGTCCCA  
69401 AGCTTTGGGA TTCTAACAA AGCTTTGTT AGTGAGAAAT TTGGAAAGAG  
69451 ATCAGGTATA GTTAAGAAC TGGGTTGGAA AGGCCACCAG GAAAGGCGAA  
69501 TATTCTGACA CAAAATTGTA TCATTTTATT TGGAAAGCATT TCAAGCCTGA  
69551 CCTGAACGAA TTGTTAGCC TCAGATACAT GCATAAAACT GTGAAAAGAG  
69601 ACATTGACTC AATTAGCTT CTTAACATG AGAAACTTTC GTGGAAAACT  
69651 AGAACTTTAC AAGCTCAGCT GGTGTTGGGG GCATCATTAT CTTGAATAGC  
69701 TCACTGGAGG AAAATGAAAT CTTAGTTGG TTCTCAGGTT TTAAAATATC  
69751 TATCATTTT GAAAAGTGTG AAGTAACAAA ATATGATCTG ATTATCTTAT  
69801 TCCTAAAATC CTTGCAGAA TTATCCCAGC CTCAATCTTC TCTTTAGTAT  
69851 TTAATGAGAA TAAGAAACTG GAAATGACTG AATTGGAAGA GTAGACTTTA  
69901 AATCCATATC TTGATGGCAT ATACATTTT CAGTTTTTT TCTAAATGAT  
69951 TAATGAGGAT TCTCAAAACT TGAGTATCTT CTATGTTCC CTTCAACATA  
70001 AAGAAATTGT ATGAAAATAT TTTAAAATT TCTAATGATT TTATAGTTAG  
70051 CTATCTGGG AATTCAATTTC TAATCATGTA CCTCATCCAA ACTCCCCACT  
70101 ATGGACAAAA ATAAAATAAA AATTATTAGT TGCATCTGAA GGCCACATTA  
70151 CAATTTCTAT GCATTATAGA AACCTGAGAA AATGTATCTT AAAAAATAAA  
70201 TGTGAACAAAC TAACCATAAT TATGAAGAAG AAAATGAAA ACTAGAAATA  
70251 AACTATTGAA AAATGTCTAT GTATCAGTTA AGTTTTTATT TTAAAATTCT  
70301 TTATGTTTAT CTCTATAATA CTATTGGAA AGAGAGAAAG GAAAACCTGA  
70351 CTTTGTTCCTC ATCCAAAGGA GGTGATTCCA CTGATTAGC CAAAATAAGA  
70401 CTTCTGGTT ATAATAAAATA ATAAAGTTT TGATTTTTT TATATGGTAC  
70451 CCCACTCACT AGGTGATCAG ACACCCCTCCT GCAAAAAAAA AAAAAATACG  
70501 TATGCAATAA AGTTAAAGTT TTATGTTATT CTTCAAGGG GAGAAACATC  
70551 TGTTTAACAC AGACCAGAAT ATTTCAACAA AGTCATCCCA ATATTTATGG  
70601 AGATCATAAA TCAAGCGAAA AAATATATTC ATCAACAACT AAACAAACTA  
70651 CATTAAATAG TCTCAAAGCA CATTTCACT TTTTCTGA CAGGAAAACA  
70701 GGTTTCACAA GTGTGGAGAC ATTTTACCAT GGCTTTAAC AGTGAGGAAG  
70751 GATGTTAAA TAAAGGGAAA AATTATATGG AAAGCTCAGA GAAAAGAGAT  
70801 GGGTGTGGCT TGAGTGACAA GGTGAGAGCA GATCTCATTAA ACTGAAATGA  
70851 GAGAGAAGGA AGGAATTTG CAAATATGGA AAGATAACTA GTGCAAGTTT  
70901 GAACAGATTA TGTCAATCAA TGTAGAATTT GGCTATCTT TTAATCAAAG

Fig. 2 (cont'd 39)

45/124

70951 AAGACTATGG AATATTTAT AGGTGTTGC TTATACTCAA AGTTTTAAAG  
 71001 AAATAACAGT ATGAATTGG TTGAACTAAT TTTTTTCATA GATAGGATTC  
 71051 TCCCAAGTTA TATAGCATAT ATATTTCTTA ACTAGTTATT CTTCCCTTTA  
 71101 CATATATTGT GCCACATTGA GTAACAACTA ACCTGCTAAT AGCTATTGGT  
 71151 TTTTAAAAGA TAATTAATAT TAGAAAGTGA TCATTTTCT GTTTCATATT  
 71201 AAACATGATA TTCTGAAAAA GCAACATTGC CTGAATGTTTC TACATTTTAT  
 71251 CTTTTGAAA ACAGGTTTTA TAAGAGATT CTTGTGAAAA GCTGAACGTT  
 71301 CTGACACTGA AATAAGTCAG CTAACCTAAA GCTAACGTTA ATTTTTGAC  
 71351 ACTGTTGGCA TGAGGTCTCA TTCCCAATT TTTCAATTAA AGCCACAGGC  
 71401 AAATGTTTA ACAGATTTA ATCCGTAGTA CAAGCATTAT TGATCTTAAA  
 71451 TTTAAGGATA AAAACCTGAT TTTAATTAGA ATTTAATATG CATTCTAGTA  
 71501 TTTACGTTGT ATAATTAATA TTTACATTCC ATGATTCCAC TATGTACCAC  
 71551 TTATTTCTTT TTGAATAAT TTCCAGTAGG AGCAGAATAA ATTTTCAGTG  
 71601 AATATTTAT TTCTGGGGG ATATTTTAA ATGGAAAATA TATTAAGTTT  
 71651 CGGTAAAATC TGTTGCTAAT TTGGCAGTGG ACAGAATATA AAAATTGGAG  
 71701 AGACTGAGTC ATTATGATGA ATTGGGTCTG ACTTTGTCA TGACACTGGA  
 71751 AATTTCCCAC AAATATTATA TTCTTCTTT ATAATAAATA TAGTCGAAAT  
 71801 GAATTGCAGT CAAGTATTG AAGACCCATC TATAAATTAA GGCGGTTACT  
 71851 GTTGATTTTT CATTATGAGA GATTCTTCCA CTCATAAGCT ACTAAAAGTA  
 71901 CATAAAGAAG GTCTGGTTGT TTGTTTAA TGTGACTGTT CTCTATCAGG  
 71951 AAAATGTCAG GTATCCGATG AAAATAGATA TATGAGGTGC CAGGTATCTA  
 72001 TTCCAAACTT GGATATCACT TCAATTAGCA TCATCTTTT TTTTTTTAA  
 72051 AGTGTCTAAG GTTACAATAG TCACCAAGATA TTCCCATGTA TGAAGCAATT  
 72101 TTCTGCAAAG GCCGCTGTGG ATGATCTTT TAAAATATAT ATTCTGGGAG  
 72151 ACATTGAGTA AAGAGAAATT ATTTACCAAGA GAATGAAGAA CCGAGGCCCG  
 72201 ATTCTTGCG TTTCTGCCAA AGATGCTGAA GGCAGTGATG AATGACAAAT  
 72251 ACATTACCAA GGAATTCTCC CTCTAAGAGG CTGACAAAGA TCTGATTTTT  
 72301 AGGATTATAT TACCAACCAAG AAGATAACCC TTGTCACTGA GCTTCTAATG  
 72351 GAAATATGGT CTATACTGAA ACAATTCTCA GTTCTTTTC TTTCTATCTT  
 72401 TTTTGAGTT ATTTTATCTT CCAAAATGA GTTATTCTG ATAAAATAAT  
 72451 TCACTTAAAT AATTATGAAA GTTCAAATT GTGCAAATAT TTTTATTGGG  
 72501 ACATCTTAAAT ATTACTCTAA ATTCAAAAAG AAAATATATG CTTTATTAA  
 72551 ATTTGATCTG TAAGCTGCTT TGTTTGTAAT TTAACATTAA TATAAAAATT  
 72601 GTATAATACA TATATTTAT TTACTTTATT CCTGTGTTGC TTTGGCTTGG  
 72651 TGAGACTAGG TCTCCACATT AGGAGTTTA CTGAATGAAA AAGTATCAGA  
 72701 ATGTAACATG ACTTTGATAT GGCATCAGAA TTTAATAAGA TGACATTAA

46 / 124

72751 TAGGAATTAG GGGTAAGTTC CAGGTTTTAC ACTTAAATAC AAATAATCAA  
 72801 TTTTGCAGGC ACAAAATACT TCAACAAAAA TCTGAAATCA TTCATTTGAC  
 72851 AAAACTTCAG GTTTCAGTT GACAATAAT ACAATACAAT GCAACAGTGC  
 72901 AATAGTGATA TCTAAATATC TAATGTAATC ATAGGTAAATA TTAGTAAGTG  
 72951 TGTTATCTGA AATGAGTGGT GTGATATCCT GCTTTACTTT GTACTGGTGA  
 73001 GTTCTGGGTG CCACCTTTGA AAGGAATAAA GACTATTCA ATCTCTTTA  
 73051 TAAGACAATA AGAAAAACAA ACAACAAAC AAACAAAAAA CCACCTCCTT  
 73101 TACTTTAGCT GAGAAAGAAG TTATTAGGTA CAGCTTGACA AGTTCAGCTA  
 73151 AGCATCCAAA TCTTCCAGGA GGTTGTTACT ACATAAAATC AACCTTTTT  
 73201 AATTCAACTA TGAGCAGGGA GATTTTATT TTCTTCGGG TACTAAAGCT  
 73251 TCCAAACTCT GTTTATTCCA CAGGAATCTG AACTTATAGA ACTAAGAGAA  
 73301 ACCATTGAAA TGCTGAAGGC TCAGAATTCT GCTGCCAGG CGGCTATTCA  
 73351 GGGAGCACTG AATGGTCCAG ACCATCCTCC CAAAGGTATA TTTAGAAATC  
 73401 ATTTCATTTTC CACCCAATAT AATAGGCATC TATTTTATT ATTAAATTACA  
 73451 GTAGAACTG C ATTACTCAG TGTCACTGTG CATTATTAAT ACATACTAGT  
 73501 TGTATTAATA GTTGTATTAA TACATACTAG TAGTATTAAT ACATACTACG  
 73551 TTGGTATTAA TGTGATCAGA ATCCTAGAAT TTTAGAACAG TGACTTCCAT  
 73601 TATCAGATAA TTTTAAACT GATCTTAAGA AATTTGGTTC TATAGTTGTA  
 73651 TACACATCTC TCTACTTGAT TCAGTGGAGA TGGAGATGGA GTGGTTGGTT  
 73701 AATACATGCA TATCTGACTT CAGGCAAAAC AAACCCATTA ATGAGTATGA  
 73751 TAATCTAGAT CTGTATTAA AAATGAAATA GTCAATATGA TGATATAGTA  
 73801 AGCAGTGGGC ATTGGGAACA ACTTTCCCTG GATGGAGGCT ATAAAAAGGT  
 73851 ACATTTCCCTG TAGATAATTT TGAAACAATA AAAACAACGG GTGAAAGGTA  
 73901 GCTCTGTTT AAATTATTCC TATGCTTAAG CAATTCTAAA CAATGAAAGG  
 73951 GGTATTTCTG CCACTGCCCT TACCCCTGGG TTCACCACTG AAGAAATGCT  
 74001 CATTATTAAT ATCGTGTCAAT TTTTTCCCTT TACATTGGTT CTATTTACTC  
 74051 ATTTCCCTGAC ACTTTCAAT GGCCTTCAGT GAGCTCAGCT CTTCCCAGC  
 74101 TTAAAAAAATC CTGTCCTAAA ACATGAATGC CTTATTATCT CTCTTTTCAT  
 74151 TTCCAGAAGA ATTCTGAGAA AAATTTATG AAGTCTTCAT ATGTCTTCAG  
 74201 CCATCTTAG ACCACTGGAG TGTAGCTCCT TTTCCCTCCA CTCCACCAAA  
 74251 ACAATGCTCT CCAGGATCAG CAGAAACTTA CATGACACTA AATTCAAGTAA  
 74301 AACGTTTATA ATTCTTATTG TATTAGACAG ACATGGAAAC AGCATTGAT  
 74351 GCTGATATTC ATTTCTTCCT ATGTGAAACA TCCGGTTTT CTAATGTCG  
 74401 TGACATCATA CATTCTTGGT TTTTCTTCTG TTCCCTTGAA ATATTTTTC  
 74451 AATATTCTT TTGTAAATTC ACTCTTTGT ATCCATTGT TAATTGTTGA

47/124

74501 TATCCTAAGC TCTCTTCCAT TATGATTCTA TGCATCCTAT TTAAAATATA  
 74551 TAGAAAATCA TCTCATACTC TAGCTGTAAT TTTTATTAAT GTGCTAATAG  
 74601 CTAATAACTG TCAAATCTAG GTCTCCAGGC CAGGGCTCTGT ATATCCAGCT  
 74651 ACCAAGAGAG AACTCCACGT GGATATCTT GGATGTCGT TTTGCATCTT  
 74701 AACACCTAACT TCTCCAAATT TGCACTTGTC TTCTGCTCTA GACCTGCTGC  
 74751 TCCTTCAGTG CTCTTTGCCT CAGTAGATAG CACCACCATC CTTCCATTAA  
 74801 GCCAGAAATC TAAGTATTCT TCATAACTCC TCCTCTCCTC ATTGAATAAA  
 74851 TTACCAAGAT CCGTTGATCC CATTCCCTAA ATATCTCTTG GATCTGTTAA  
 74901 CTTTTCTCTG ATTTTACTCT TGCCATCCAT CACCTCTCTC CTGAACCATG  
 74951 ACCACAAACC CCTAAATAGC CTTCCCTCTC TTAATCTTAT CCTGCTTTAC  
 75001 ACCAGTCCTC ACGCTGAAGC CAGAATAGTC ATTAAGAAAC ACATCTACAG  
 75051 GTATCCCATT CATTGCCTT AGAATGGAAT ACAGACTCCT CAGCATGACA  
 75101 TAATCTCTCT TCACCCAGCTT CATTATTCA ACAAAATATTT ATTCTATAACC  
 75151 AATTAAGTGC CAGATGATGC ACATATAGAC TTCTTGTCT GTTGTGTCAT  
 75201 TGCATATTCC ATATTCAGC TATCCTGAAT TGTTTCAAT TATTCTATAAG  
 75251 TTCTTTATGA ATTGTGTTCA TTCCATTGG AATATTCTAC CTTGTTGAT  
 75301 CAGCATAAAAG ACTTTTCGAG ACACTGCAGC AGCAGTGAAC CTAAATATGT  
 75351 TTCCATTGACC CCTACATTGA ATGACACCCCC CTGTGATATG TTTCTGGAAG  
 75401 CAGCAATACT TCCCTTCTTA AAATTACATT ATACTTGGG GCTTTTATTT  
 75451 AAGGTATGTC TTTCCTGATT TACAATAGTA GAGCTTGTCTT TTTCACCCCT  
 75501 TTGAAAGACA TCAAGATGCC CATGATGATG TCTTGATGT AACAGGGTT  
 75551 TATTTGAATT TTTAAAAGAA GAATAAGTA ATTTTTAAAT GAATTTCAT  
 75601 TTAAATTTA GGAAAACAAT TATATAAAAGT GAGATATGCT TAAATTGAAG  
 75651 GACAAAGTAG TTCTGTAGGG GCTACTCTT TCAAGACTTT AGCAACTTTC  
 75701 CATGTGGGGG AGTGATTTAT GTGATGCATG GAAAATTACT GCATATTTAA  
 75751 AGCTTATCTT AGAGCTATAA TAAAGCAGCT TATGTTCTAA ATCTTCATGT  
 75801 CGTAAATAGG TCCAGAAGGG ATTTAAAAG CCTTAATCCT TACTTTAAC  
 75851 CAGCACAAAGT CACTGAAGTG AAACCTGCTG AAAGGATTCC TTTTATGTTA  
 75901 GGCAACAGGT AGCTGAATAT ATCTACAGAA ATTGAAAAAT TGGAATTCTT  
 75951 TTGCTCAGAA ATGTGGGAGG GGTGGAGCTT AAGTAAAAAA ATAACAGTTA  
 76001 ATATCTAAAT TGATCAAGAA ATATGAAAAA ATAATTGCT AGGTTTAA  
 76051 ACTAACAAAA ACCATGGTTA TAAAGGTTG AATATATATA GGATAGTTAG  
 76101 ATTGTATTTG TGTAATATTA AAACTCAGCA TTAAATTTAA TGAACACAAA  
 76151 GTGATTCTTA TCACATTGAC CATTGACATT ACATGGAAAA AATAGTCAGT  
 76201 TGGACTAATT ATGTGTCTTT CCATGGGTTA TTAAGGTAAT TGTATGGCAT  
 76251 ATAAATTTAT ACTGGAAATC ACATTGAAAT TCACTTTAG AGGCCCTAA

Fig. 2 (cont'd 42)

76301 AATATTTCTG TAATATATAT TTTAACATA TGATCTAAA AGATATATTT  
 76351 GGAATGACAC AACAGTTTA TAGACAGGCC TGACTATCAC ACAACCACAC  
 76401 ACCAATTGT GAATGTGTT CTATTCCTC TAAATTAATG CATCACATTC  
 76451 ATTAACAAAG TTTGATAAAT GACTATAGTC TATAATAAAA TATTTTGTT  
 76501 TACAAACATA TTTAACACC TGCTATTAAG TATAGGCATT ATCAGATCTT  
 76551 AAAATACAAA GATTTAAAAA ATTACCCGTG GGTCAATGGAG CTCACAATCC  
 76601 ACTGAAAAAA TAATGTTGT GATAAGAAAT TTGAAAGTTG AAGGTAATAG  
 76651 AAAATTTAC CTTTATTTT CAAAATGTAC CATTGCTTTC TAAGTCACTA  
 76701 CTTCTGTGTA AATATGGAAT TGTTTTCTC TAAGATATAC CAAATATAGT  
 76751 TGGATAACGC ATGTATTAAA ATTCTGTCAG CACTAAGTTG TTTTTAGAC  
 76801 ATAGTGATAG GCAAACATAG TTATATTGAA TGAAAATTA GAATCAAATT  
 76851 TATTAAACAC TGTGTACTGA TTGATACCAAC ATGCCATATG CTTGTATAGC  
 76901 AATACAAGGT TTGGAATTAA TAATGGTAAA CAAAATAGAT ACGGTCTTG  
 76951 TCTCCATAGA ACTTTAGTC TAGTGGGAGA GCAGAAGGTA AAGGAATGTA  
 77001 TGTGATCATT GGTGAAGCTG AACATGTATA CCCAACAGT TATAAGTTCC  
 77051 AAGATGGACA ATAATGGGTG CCATAGGGAA GGAGGGTACC AAGGAACCTA  
 77101 CTGGAGGTTA CATAGGGAAAG ATTATTCCAA GGTAGTAATA TTTAAGTGAA  
 77151 TATCCAAGGA ATAATTGTCA ATCACTTTAT AAGTACTGAG GGAGGAGTAT  
 77201 TTCAAAAGAG CTTTGAGGCG GAAAATAAAT TAGTTCTTT ATGGAACTAA  
 77251 TGTAAGGAAA ATACTAAGCA AACATGTAAT AAGAAGAACAA CGGTTGATGA  
 77301 GTTAAGAACT GACAAGATTA CTGAAGGATT GTAGGCCATA TTTAGAAGTT  
 77351 GGATTTTTA TCTATTCTTA TTAAAGTGTGAG AAGTTATTGA AAGGTCTTAA  
 77401 GTGGGGGAGT GATGATGAAG TTTGCCTTTT AAAAAAGATT TTTCTAGCTA  
 77451 TTGTTTATAG AATGGTTGA AGATGAATAA GTCCAATAGC TATACTTGCT  
 77501 GTAAAGGTTA TGTGTTGAGC TTGAAGCTGG GCAGTGGTGA CACAGAGGAT  
 77551 GGGAGATGGA AAATGACGAG TGAACAAACA CATACTGAA AATTAAAGTT  
 77601 TAAAATAGA CCTCTCCATT AATTCAAGATT GCTGATATTG ATTCGGTTAG  
 77651 CCATTCTTTA CTGAACCTTA TGATGCCCA TATACTGAAT TAAATACTTA  
 77701 CAAGCACTAA AAAAGAAATT GTTAGGGAAC AGTAAAATGC ATTCCTTCA  
 77751 TTTCCACAATA TTATTAATAT TATGGCTTGT CTAATCTTTA TTGGTGAATG  
 77801 CAGTCATAAT TGAAGGTAAC TGATACTTCC AAGGACTACT TTTGACCTAG  
 77851 GATTACTATC TTTTAAAAA TTTAGTATTA AAGAAGTCAA ACACAATTAA  
 77901 TTAATTCTGG ATATAATAAA AATTCTGAA TACTTTAATA CTTTGTGCTT  
 77951 TTCTATTTGT GAAAGTTAAT TATTAGGAAC GAGCTAGCAA ATGCTACTTC  
 78001 TTTTCACAAA AGCTAATGGC CAATCACAGC AAAAATTAA ACCACTAAGA

09/914549

49/124

78051 AATACTACA CATATTCTTC TATTGCCAT TTATATGACT TCCATAATAG  
 78101 TTGATTAAG GATACCGGAT TCCTTTATTG TTGAATTAAA ACCTCCTACA  
 78151 TGAAACCTT GATTTAGGTT TAGAAGTTGG TAATGTTTG GCATGCAAAA  
 78201 CCAGTTAATG TTCTCATCAT TACTTTTAA ACAATGTTA AGAGATGAAT  
 78251 TCTAGGGATT ATAAAAAAA AAAAGCTGTA TGTGTTCTT CCTATAAAAT  
 78301 TTTTCAGCAT GATTGCCTCA GTAGAAAAAT TAAGGGACTT ATTGATATAT  
 78351 ATGTATATGA AGGTGAGGAT ACACATATAC ACACACACAT ATATATGTA  
 78401 GTAAATACAT ATATTACATG TCTATCAATC CATACTACT CATTATTATT  
 78451 ACGTTTGAA AGCAACCAGT TATAGTTTG TTGCCATGGA TCATTTTAC  
 78501 TATTCAAGTAA ATCAGTCAAT TGAAGAGGCT TGATTTATG GTATTAGTTT  
 78551 TTTGGAAACT GTCAGCTTA TAGTAAATT TGACATCTTA CAACTCCAC  
 78601 TGAGATTTT TTGCTTGACT AATCTGCCTT GATGCCAATA AGTATATTAA  
 78651 CGGAAATGGA CTAAAAGCAA ATGTGACTTG AAGCACAATT TTGTAATTT  
 78701 TCTTAGTGTC TCAGTAATAC TTAATACTAG TGCACTTTAG GTAGGAAAAT  
 78751 TTTCAAGTTG TTTTATTCTTA AATAACTATA AATCTTATAG TTGCTTGTAT  
 78801 AAAAGAAACA GATACCTTA ACATGATTAA ATATCAAATG CTATTCTCTT  
 78851 CAAAATATCT TAACTAAAGA AGCACTGCCT GCTCTTAGAA GTTAAGCAAG  
 78901 GCCATACCAT ATGCTGCGTA CATGGCTTT AACACAATGG ATATTAGAAA  
 78951 CAGCCTAAGG CTGAGCCTGG CTCCACTATT TTTCAGCTAT GTGACCATGT  
 79001 GAAAGTTACA TTTAGTAATT AAACTCATT CAGTAGTTG CTTAAGAAT  
 79051 AAAATTAGGT ACTCCGGGG CATATCAAGC ATATTGTAAA ACCTAGTTG  
 79101 ATTATTATTT GTTATTGGTA TTACTATTAC TATTCTATAA TAAGTCATGG  
 79151 GCAGGCAGTA GGGGTACATT GGAAGAATTG CACTGTCTTA AATATGCTCT  
 79201 CTGTTTAACT CACAAACTCA GTCTACCTAG GCTTTCTTG GAGGATCTGC  
 79251 CTTTCATTGG CTGTTGACT TTGGCCAAGT TACTTAACCTT CTTTCACTT  
 79301 CAGTTCCCTC ATCTGTGAGA TTATGTGCTT ACATGACTTC AGGTTTGTT  
 79351 TTGGCTCTAA TATGGTATGA TTCTATGAAA TGGAAAGTTA ATACATTGG  
 79401 CTCTAGTAAC TGTATTTGAA GCACAAATAT TAAAAGCAC AATTAATTCT  
 79451 CATTCTGAGT TTCCATTTAC TCTTTAAAT TAATCATTCA GAATAAATCA  
 79501 TTTTGGAAAGA GCTGCTTGAT CCAGGTATTC AGTAGAAATC ACTAGCATAG  
 79551 CATTAAATT TAGACAAAC TGAGAACTCA TTAAACTGCC AGGGCTATGG  
 79601 ACTTATATGA GATTCTCATT AAATCTTAAT GTAGATAACT CAGTTAATTA  
 79651 AAACAAATAT GGTTGACTT TATTAAACTT CTAAAGTCAA AACTGCATTG  
 79701 AAATTATCTG TACAAAGCCT TGTTGACCTT TATTAGAGAA CTGCCTCTCA  
 79751 AAAGACCTAA AAGACTTATT TGTTCAGATC GAGACTCTTC ATGAGCCAAT  
 79801 GTGATACTCT CCCTCTATTG CTAGATCTTC GCATCAGAAG ACAGCATTCC

50/124

79851 TCTGAAAGTG TTTCTAGTAT CAACAGTGCC ACAAGCCATT CCAGTATTGG  
 79901 CAGTGGTAAT GATGCCGACT CCAAGAAGAA GAAAAAGAAA AACTGGGTAA  
 79951 GTTACCATCC TTCATCTAAT TCAGAAGCTT ATTAATGCAT AATGTGTTAG  
 80001 GCCTTTCT TTGGGGCTTT AGTGATCTGC AGTAGTTAC AAAGGGTCCC  
 80051 ATTCAAGCTA CTGAGACCTC AAATGCTGCA CTCATCACCA AAATTGGAGT  
 80101 GGCATGTACT GAAAAGCATA CATTAAATG TTGGGACTAA ACTTGGGTTT  
 80151 GAATCACAC TATATCTAGA CCTTTTGAGG GGCTGAATT TTCTAACCAA  
 80201 TAAAAAGACA GTTAATAGCA ACTATATTAA TTTGTGAATA TCATTTATTC  
 80251 ACAGATGTTA TCTAATTTT CTATAGTATA ACTATACAAA CTATGTAGTA  
 80301 TAACTATAGA GTTATACTAA AGAAAAATAA GATAACATCT GTGAATAAAAT  
 80351 GGCTTAAAT AGGGGTTTAT TGTGGCATA GAGATGAAGG AAAAGTGAAA  
 80401 AAATGATGAT GATGGTGATG ATGATGGTGA TAGTGGTCTT GGAGGAAAAG  
 80451 GAGAATGGGA GTTAATAAAG GGAAAGAATA AACAAATGAAA CTCTCATTCC  
 80501 ACCTTTGGAA TCGACAGGGC TTACCGTGTG AATAGTTCA CCCTAAAAGA  
 80551 AATCAACCAC ATTAGTGTCT GCTTGATGTT TTTAACCAAG AGAATATAGC  
 80601 AGAAATATAG AAATGCACCT TAACAGAACT GTACCTTAAG TTTGCTAGTG  
 80651 ATATAATTAA TGATATTGAT CAATAGCTAA ATAGCCCAGG GGAAGATACT  
 80701 GTTACTGCGA AAAATTAAA ACAATGGAG TCAATGATTT CTTTTAACAC  
 80751 CAAAAAAAT ATGTAGATTT TGAGTAAATA CAACTCTTGA TGAAATCCAG  
 80801 ACATAATTAT CAGAGGATTT TACTGGAGTG CTTTCTACAA ATAATGAAAG  
 80851 AAATATCTT TTATCTTAAA AAATGTTTAT ACAGGTAATA TTTTAAAATA  
 80901 CTGATCAGCC TTCATTCCCT TGATTTGTAA TTCCACACTC TTTCATGTTT  
 80951 CTGCAAGGTG AACTCTAGAG GAAGTGAGGT GAAXATAAAC CGTGGACAAT  
 81001 TTGGCATGGA TXTATAAAA AACCTACCT TGGCATGAAT GCTATCCATT  
 81051 TTGGCAGTAG GCTTTATAC CTTTAAAC AGATTACCTT GTATGTCTT  
 81101 TCTTTGTGTC TTTTCATTAA AATCTCAAAT TTTAAAGAGA TGAAAACCA  
 81151 CTTTCTGAAT AGAGCTGTAG GGGATACCAA TTCTGGTTT GAGTAGTCTG  
 81201 GGGTTGGAAA ATTTGAATAG AAAAATCACA ATTAATGAAG TGTTAGGTGA  
 81251 ATTTGATTTC ATTTGCTTT TTAAGTTGT ACTGTCAGCA GGACATGACT  
 81301 TGATTGTAGC GCTAAAGTGG CCATTTAAA CAAATTGCCT TGAAGAGAGA  
 81351 AGCATTGGGA ATGGAGATC

Fig. 2 (cont'd 45)

## Human genomic sequence

1 GAATTCCCTGG TGGAGAACAG CACATGTACA GATGGGGTGA AACAGCATA  
 51 CGTACAGGTA GGGGTAAGCT GGTGCTATAT GAGAAAGCAT CGAATAAGTT  
 101 ATTAAGTTG ACCTGCTTGG GAACTGAGGG GCAGGTGTGA GGGATGAAGC  
 151 AGGAGTAGGT AGGGGCTAGA TCACAAAAGA TCTATGCCAG TCTTCTCAC  
 201 AGTGTGATTC CCAGCCCAGT AGCATGATAT CACTGGGAT CTTGTTAGAA  
 251 ATACAAATTC TTATACATCA CCCTGGACTA GACCACCTGA ATAAGAAAAG  
 301 TTGGGCATGA GCCCTACAAA TTTTTAAAAA AGTCATACAG GTGATTGCAA  
 351 TGCATGCTAA AGTTGAGAA ACACTCTTG CTGTGGTTG AATATTTGTG  
 401 TCCTTCCAAA ATTCAATGTAG AAACCATCTC CAATGTTATA GTATTAAGAG  
 451 GAGGGACCCCT TGGGAGCTGA TCAGATCATG AAGTCTCCTT TCTTATAAAAG  
 501 GGGATTAAAAA GCCTTGGCCC TTTTACCCCTT TGTCCATGTA AGGACACAGT  
 551 GTTGGAAAGCA GGGACTGGGT TCTCACCAGA AACAGAACCT GCCAGCCTCT  
 601 TGGTCTTGGGA CTTCTCAGCC TCCACAATTG TGAGAAATAA GTTCTGTTG  
 651 TTTATAAGTT AACCAAGTCTC AGGTATTGTA TAATGGCAGC ACAAAGGGGC  
 701 TAAGAAAATG TTCTATGCCCA TAACAAGAAA TGTGGTCACT TTCCTGAAGG  
 751 AAATGGGGAT ATATATAAAG ATGTTATATA AGACTCGTAA TATTTATTTG  
 801 GAAGGCTTGC TCTGCAAGCA AGGTGGAAGA GCAACATGAA GGAAGCGTGG  
 851 TGGAGGTGAG AGGACTGGAG GTTAAGTTGG TAGGGAGATA CAGGAAAGAA  
 901 GCTTATGACA CTTGAGTTAA AATGTAGCAT CCTTCTTATG TGTAGGGCTC  
 951 ATAAAAATGT ATAGTCTAAG ATAGAACACA GAATACTCTA TGAATCCTGC  
 1001 CCACAAGGTG TTGGTAATCT AGATTCACCTT TTTTTTCTG ATAATGCCAT  
 1051 CCATATGTAT GGAGCGTCTA CTACTGTATG CCAGAGTGAC TCTGGAATCG  
 1101 GTTTGGTTGA TCTAGACAAG ACCATAAGGA GAGTCCCCTT ACTACCTCTT  
 1151 CTCCAGGGGA GGGATTCAAG TTGAACCTAGT ACTTCAGAGA CTGTTTAGTA  
 1201 ATATCATGCA TGAAAGGTGA TGGTTAGGAC AGAAAAATAA ATGGATTGCA  
 1251 TCATAATTCC TCAGGTTCTC CAAATATGTG GTGGCTCAA ACCATGTGAA  
 1301 TTGGTCTGCA CATCCTGTTT GGGTTGCGTG TCAGCAGTTG AGATCTGAGC  
 1351 CTTATTTGTA ACAGTGAAAC AGTGAGAGAC CTGCCCTCTA AGAGCTGTTT  
 1401 TTCAGCTAGG AATAGAAAAG GGCCAGGCTA GACTCCTCTT TCTGCTGGAT  
 1451 CTTGCTTCTT CTCAGCAATA GAAGTAGACC TGCCTCCCTA GCTGTAGAGA  
 1501 AAAGGTGCCG GTAGGCGGGC AGGTGAGCCT GTGGATAATC CTGGAGTAAA  
 1551 GGTTCAATAG ACCTTCAAGT CTATCCTACA GGATTCGGAG TGAGGGGAGA  
 1601 GAAAAGGAGA CGCTTCTCTG GCTGAGAGAG GAAGAGAAAA AAAAATCCCA  
 1651 GATATCTGAC AGCTATATCT TCCCACATCACC ACCTTCTCT AAACCCATGC  
 1701 CTCTCTGTTT AGTAGGACAT AAAATGAAGA GTGACCCACC CCCCCACCCCC

1751 AGCCCATCCC CCGTTTGTAG GTGTGCTTTC AATGAAAATA AGTCGGTGT  
 1801 CATGGACGGA AACTAGAGCA GCTGAAAATA GATGCAAGAC TTGTTGAGCA  
 1851 TACAAATCAT TTCCCCCTTA GTCTCCAAGG GAGGAAAAAA AATCCCTCTT  
 1901 ACTCTCCTTG CAGCCTGTGT TCTGCATTCT GGAGAGGAAG CTGAGGCTGG  
 1951 TCCTCAGGCG CTCCTCCCGC CGTTCCCGCA GGAAAACCTTT CTGCAGGGC  
 2001 CCGCTCCGTC CATCCCGCGC GGTTCCAAGA CGGTGGGCCT CCCGTGGGCT  
 2051 CCTCTCTTGG GCAAGGGCCC AGACCCCGCG ACCGCCCTGT CTCTTTAAAT  
 2101 TCCAGCTGCG CGGCTGGAA ACAGCGCCAC TCGCCGCCA GGCCGGCTGG  
 2151 AGGCTGAAGA GCGAGCTCGC GCTTTCGCTC CCGGCTGCGC GCCGCGGAGA  
 2201 GCTGGGCTCG GCCCGCGGGC TGCTAGGTGG CGGCGGCGCG GGGCGGGAG  
 2251 GCGCGGCCCG CGGGAGGAGG GAAGAAAGAG CGAGCCGGGC CGGGAGAGGC  
 2301 GCCGCGCCCG GTCCCGCGCC CGGTCCCGCA CCCGCTCTCA GCGGCCAAG  
 2351 CAGTTTCTTT CTGGGTGACA AGAATGTGCC TCGGTTGGTT TTTCTTTTT  
 2401 TTCTCCATCT CCTTAAGACG ATTTCCATAG TAACCTGATC AAGTGGCTCA  
 2451 AAATCGCAA CCTGAGGATT TCCGCGGCC GCGGCAAGA CCTCGGCCAG  
 2501 GTAACGCTGC GATCTCCTCC TCTTCCATTG CAAACCGCTG CGCTCCTTGC  
 2551 AAAGTTCTT TTGTGGAAA TCGCCCAGCC CAAGGGAGCC CGGGGTATTT  
 2601 GCAACAGCGT GTTCATTTCAG AGGTGCCTGT CACGGGTCTC CTCCCTGCTG  
 2651 CTTCTCCAGG ACCCATGATG AGATTATTAA TAAAAATTGT TTTTGGTCGT  
 2701 CTCCCCCGCC CCCTCCCCCTT CTTTATTTTT TTCTCTTCG CTGCACTCTT  
 2751 CTCGGCTTTT CCCCTGACAC TACTGATGGG GGTGGGGGG GACGTCGGGG  
 2801 ATGGGGGTGG CCAGCGCGGT CCTGGGAGTG GCGGGTTCGG ATGGGCTGGC  
 2851 TGCCTGGGC CACTTGGGC ATCTCGCGT GGCCTGCGCC GGGTCACGG  
 2901 GGAGGGCTGT CAGCGCCAGG GCGGCGGAAC CCGAGGTCTC CAGACGAGTG  
 2951 AGGGAGGGAT GCAGGCTTGG GGGTGTGAA GCGCTGGCT GGTGGCTGGT  
 3001 GAGCGTCAT ACATCATAGC TCTCCTTCCC ACTCCCCCGC CCCTCTTCGG  
 3051 GATTCTCTCT TTCTCTTCCC CCGTCCTCAT TTCTTCTTC CTTTACTCAC  
 3101 CACTCGCTTC ATTCTCTTCC TTCCATTTC TCTTTTTTC TCCCCTCATT  
 3151 TCCTTTTTT CCTTCCCTT TTAAAGAAAG GGGAAATCGTT TGTAACCCTT  
 3201 TCGTTCTACC AACGTGGAAT AGCTGTGAAA CCTGCAGCGT GGTCACCTCA  
 3251 GCCTGGTCGT TTTCAGACCC GTCCTCATCC ATCAACATAT TTGTTTCCCG  
 3301 AGTCTATTGA TCTCCCTGAA TTCTACAGAA ATGCATTCTA AGCTAGGCC  
 3351 CTGTATGTCA GAATCAGTTC TGCAGGTAGC TTCCGTGCTC CAAGTATGAC  
 3401 ATGTATTGTA AGGGCTGCAT CTGTTTAAA CCCACATAAG CCATGGGTAT  
 3451 AAATAAATGT AGCTTGAAA AAAATCTGG CCTTATTCTA GATAAACTTC

Fig. 3 (cont'd 1)

09/914549

53 / 124

3501 CCTCTTAAAT TACTGATATA CTCTTCTCCC TCTTGACAT TTAATTAG  
3551 GAAAGTTGGG AGACAGGTTC TTGTCCTCCA GTTTTAAGG AGCAGGCAAC  
3601 TTCTATTATC TTAATTTCT CGTCTTGAA CATCACTCAC GTTTGCAC  
3651 CCCAGTCAGT GGAACGAGTG GGTCATAATT AA

Fig. 3 (cont'd 2)

## Human genomic sequence

1 CCTGCATTAT TGTTTTATC TGACTTCAA TTTGGTGT CCCTGGTGG  
 51 GTGGGTTTC CTGACACATT TACAAGATGC TTTGGCAGG TTGGCTGGAA  
 101 TTTGAAGGCA CATTAAATTG TAGGTGCAAT AAAATATTCA TTTCTCTTG  
 151 TTCTGGTTT GAGATGTCAT GCCCTTTGG TCACCTATAT TTTGGTGTGA  
 201 CTGTGTGT GTGTGTATGT GTTTGTGTGA AGGATTTAAC AAAGTCTGTT  
 251 CTAACTGTCA TGTGATTGAA AGTTAAAAGG TATGTTAGTG ACAAGCCACA  
 301 AATTCTCTT ATTTATAGTA CATTGATCCT GAAACCATT TTTCCCTTGT  
 351 GATTCTTCT GTGCATGGAT CATTAAACGA AAGGTTGGCA ATGATGAGCT  
 401 ATTTTTTAT AATAGAAAA AAATCCTCA AGTTACTTA CCAAGTCATA  
 451 TTTTATACA GAGGGATTAG CAAATATTG TGATCTAATA TTTTAATAGA  
 501 CTGAATTGCT GACCACTGCT AATTACCAAG AATATATTCTT CTTAATTCTG  
 551 AAATTGCTGT ACCTCTCAAG TTGTCGGAG GACTCCAAGT GACCCAACCT  
 601 GTAACCATG GCAACAGGAA GTGGTTGTTG TGGGTGCAAG CTGAAGTGTG  
 651 CACATGGACC CGTACTTTGT TAGCACTCGG GGACTTGATA TGGAAAGAAT  
 701 TAATGTACTG GCTTTTTGT ATAGATGAAT GTTAACCTTC TGACATTAGT  
 751 CAGAACTACA TCTCCCAAGC CTTGTTTGC AGTGTCTGTC CCTTGCTCT  
 801 TCACTTACAG TAAGTCCTTA CTTAACTGAC TTGATAGGTT CTTGGAAACT  
 851 GCAACTTAA GCAAAAGGAA GTATAATGAA ACACTTTAT CACAGGCTAA  
 901 TTGGTAGAAA CAAGACTTAA GTTCCCAGG CATATTCTG GTCACAAAAA  
 951 CATTCCAAA CTTCTCAAAA CACTCAATA TTAAGCATTG AAATACATGT  
 1001 AAACATATGTA TATATGTAAG AAAGGTTACT ATAAACCAGA TCAATATTAA  
 1051 CCCAATTATT TAAGTCAGG GTCTTAGGTG GCTGGAGCCT ATCCGAGTAG  
 1101 CTCAGGGCAC AAGGCGGGAA CCAGCCCTAG ACAGGACACC ATCCTGTTGC  
 1151 AGGGCACGTT CACACATGCC CACACGCAGG CTGGGACCAT TTACATGTGC  
 1201 CAATTCACCT ACCATGCACA TCTTGAGAC GTGGCAGGAA GCAAGAGTAC  
 1251 CTGGAGAAA TCCATACAGA TATGGGGAGA ATGTACAAAC TCCACCCAGA  
 1301 CAGTGGACCC AGCCAGGAAT CAACATTGG GCAACATTAT AATGAAACGA  
 1351 AGTTGAATGA AATGATGTCG TTCCACGACC TGCTGTACTT GAGGGGTGTT  
 1401 ATAAAATTCT CAGAAGACAG AGGTTAACG CTATTTTTT AATAGAAAAT  
 1451 AACTTATAGA GAAGTGTGCA CATGTGACTT TGTGTGTAGC AGGAATCATT  
 1501 AGGATGAGAA TCAGACGTAA GAGGTGGTGC CAACATGAGG AATGTTGAGA  
 1551 TTCAGGGAGC TGTGGATGGA AGTAGAAGCC AGAAGGCCAG GGTTAGGTTC  
 1601 CTACTTCTTA CTGTTTCAGT TATTGCAGTG TTGGCCTGTT TATTCACAGA  
 1651 TGTCACCTAG CTTTGTGTTTC TCAAGAAGAA AAATGAGCAT AATCTTCCT  
 1701 GTTATGAATT CTTAACACAA CAGGACATAA CCACAGACAC AGAGGTGCAC

Fig. 4

09/03/14 09:06 09/03/14  
09/914549

55 / 124

1751 ATATGTAGCA GTAATGGATA CTAAATGATA CACTCGGAGG AAACAGAAAA  
1801 GACTTCTGAA TAGAGACTGG AGATACTTCC TTGGACCATT GATGAATGGG  
1851 CAATGATGCA TTTTGTCCTT CCATTCAGAA GGCTAATATA TTGCTCTCTA  
1901 TGTTCTATGG ATAAAGGCAG TATATGCTCA AGGATGAATC ACATAATATG  
1951 CATAATAAAT CCAGCAAGCA TTACCCCTTT ACTTATGTGA CTGCAAGTAG  
2001 GAATACATTT CCCCCACTCT AACCATGTAA GATTTCTTTC CCTTCTCCCA  
2051 TTTTGTAAAGC AAAAGTAAGT TCCTGAAAGG TTAAATGGAC CTCAGGATGG  
2101 GAAAAATCCC CAGAGCTATC TTTCTGCACA GACTTCATTT TTTCTCCCAA  
2151 GTCTGACTGT CAACTGCGAT ATCTGATATG AGGCTCTGGT GCTGATGTTT  
2201 CCATAGGTCA TCATCCTTCG GTGTCCCAGA TGAAGTCTCA GGTCGAACAT  
2251 TGCAATAGCA CAGATTCTGA ATTTAATGCA TCATTAAGT TGGTTATGTA  
2301 ACCCAATGGC CTTGTTAACAC TCCAGATTT TAAAATTATA TGTATTTACT  
2351 ATTCTCTTAT TTTAGAATGA TCTCACAAATG TTCACAAAGAA ATAAGCCCAG  
2401 TCCCTGCAAA GACTTTAAAA GCTGCTTGTGTT CACATCATTA GATTGTACAA  
2451 CGCTTGTACA ATGACACTTT TTGCTAATCT ATGCAACATT TTTGTAACAA  
2501 TTGTCACAT TTTAACTACT TCAGATAATC AGGACCTAGA GACTTCAAGA  
2551 TCTGGAAGCA TTGCTGGTGA CATAGAGCAA AAACCTTCTT GAGAATAGGA  
2601 AGTCAGTGTT TTGACAAGTG ATTTATAACA GTTCAGGTAT AGCCAGGAAG  
2651 GTTTGAAACA AACCTTAAGT ATTATTTCTT TCATCTTGAT TAGTATATAT  
2701 TTATATGTGA TCTATTATG TATATTAATA GATTTTGGG TCTTATAGCC  
2751 AGCTTTCATT TTTCTCTATT GGAAAAGATC TAAGTCCCCA TCCTTCCTTG  
2801 GTGGCTTTG GTAGGTTGT AGACAAAACA TTGAAGAATC AATGGTACCT  
2851 TTTATACATT AATACTGCCA ATATGACCAT AAAATCATAT TTTTTGGGAA  
2901 TTTATTCCCC CGATCAAAAG AAGCATTGT TATTGAACAC AGTCTTATGC  
2951 TACCTTATTA AGATGTATCA AACACCCCTGA TTGATCAAAA ACACCTCAGT  
3001 CCATTTAAG GCAGTATTGC CCAGCAATTAA AAGATGTAGC TTCTGGAGGA  
3051 GTCTTCTGA GTTGAAATTC AGTACTCTTC CACGTACTAT ATAGGTGATC  
3101 TTGGGTAAAC TTCTTGAGTC TCAGTATCCC CATCTGTAAA ATTGTTGTAG  
3151 AGAAGAATTG TTGTGATGAT TAGGTGAGAG AATATATTAA TGTAATATTT  
3201 AGGAGAGCAA CCAGCATGTA GCATATATTC ATTACATATC AATTCTATA  
3251 TTATTGATGT TCATACTGCT GATGTTGAAA TGCACAGGAA GGCCACAGTT  
3301 ATTTTCTGTT TAGATTGATT TTTCTTTTAA AGTCTGAACA TAAACTGTAA  
3351 TACTGTGCTT ATTTATGTAG GAACTGTGAT CTCGTCTCCT CCTTTTCCCA  
3401 TCTCCCCCTC TCTACCTTAG TTTTCCTTA TAGTCTCAAG CTGAAAACAA  
3451 TGACCAGGTG CCTAAGAGAT AAGAATACTC TTTCTTTGA ACTCATGGCA

Fig. 4 (cont'd 1)

09/914549

56/124

3501 TTAGCAGTGA CCTGGATGAG ATTGGAGGCT ATTATTCTAA GTGAAATAGC  
3551 TCAGGAATGG AAAACCAAGC ATTGTATGTT CTTACTTATA AGTGGGAGCT  
3601 AAGCTATGAG GATACAAAGG CATAAGAATG ACACAACAGA CTTTGGAGAC  
3651 TTGGGGAAAG GGTGGGAAGG GGGTGAGGGA TAAAAGACTA CAAATAGGGT  
3701 GCAGTGTATA CTGCTTGGGT GGTGGGTGCA CCAAATCTC ACAAATCACC  
3751 ACCAAAGAAC TTACTCATGT AACCAAACAC CACCTGTTCC CCAGTAACCT  
3801 ATGGATATAA AAAAATTAAA AAAAGAAAA AAAGAAAAT CTTTTTGCA  
3851 GGGGGCAGGT AAAGGGTAAG AGGGCATCCC ATTTTGAGT TTCTAGAAA  
3901 GCTT

Fig. 4 (cont'd 2)

## Human genomic sequence

1 CTGCAGGAAG CAGCAGCAAG GTCCAGGGAG CCTCTAAATT AAATAGGAGA  
 51 AGTCAGAGCT TTAACAGCAT TGACAAAAAC AAGCCTCCAA ATTATGCAA  
 101 TGGAAACGAA AAAGGTAAGT GTTGTTACA TCATTATGAC ACAAGTCCAA  
 151 CATGAGTCTT GTGAATTGCA TGCTAAATCT AATATTGAG CAGCGTAACA  
 201 ACTTTGGGCC TAGAGATGTT ATCAGTGGAG TTTCTTATG TTTCTTAAC  
 251 GTCCCCTCCT GACTGCCAGC TTTCTTATCT GAAGAACATT TTAAACAAAT  
 301 AAACTCATTC ATTTAAAGT AGTTAGTTAT ATATGCAAGT ACAAAATACTG  
 351 TTTCTCAAAA ACAGGTCTT CCAAATGCAT GTAAATCACA TTTCTTATG  
 401 TCTTTTTATG TTTTGAAAAA TGTATCCTGA AATCATAAAG CCATATTGAA  
 451 TTTATCTGAA TCCTTAACCT CAGTTAAGGT AAGAGCCATA AGTGTTTTG  
 501 ACAATTAAGG TTGGAGCATE AAAATTGAA ACATAATTAC AGTAGGTTT  
 551 TATCTTGCA AGCAGCAGAT CCCAGAGATA TTATGACCTC AGTTTCCCC  
 601 AAAAGACAAA TTATTCATAT TTGTTTGT TTCTTGAATT AGTGCATAAT  
 651 ATAAATATCA AATCACAAAA TCAAGGACAT TAAATGAAAG TGTCTGTTAA  
 701 AGGCATATTA TAAATGAATC ATAAGCCACA CAGTTCTCTG TGATGTACGA  
 751 AGTGGGCATT TAAAGAGGTG CTGATTTGAT GCTTGTCACT GAGTAGCAGA  
 801 GAGGACGGGG ATGAGTATGT GTAGTTACA CCTCAATCAT GAGGAAGTCA  
 851 AGAACTTGTG CTGTTATAAG TAGTATGGCT GTGTGAGGAA CTAGGGTGT  
 901 CTGCTGGATT TTGAGGAAGT ATTTTCAAAT CAATAGAACT TCAAACTTT  
 951 CTTCAGAGTG TTGGGCTCTA CATGGAAAAA CACATGAAAT TAAAAAGTGG  
 1001 CACAAATGTT TAGTTAGTAG AACATCTGGC TAATTGGAT CAAATAATT  
 1051 AACCATGTGG GAACGTTTT GCTCAAATA GATAATTGTG AATTGTTCA  
 1101 TATAGGCAAA TGATTAGACA ACTTCCTCTT CCTCAAATGT GAACGGACAG  
 1151 ATGTGATCTA GAAGCAAGAC ACTCTTTGT GTAAATATTG CCTTTGGCCT  
 1201 AAAGCAAAAG TGGACAGACT TTAAACACCT GAGAGCAGAG CAGTGTGT  
 1251 TAAGATTGCA ATATCTTAAG CTCTTGAGTT AAATGGAAAA TGAAAAACAA  
 1301 AAGTGTATAT TTGGAAGTTA GGAATGTTT CTTTAAATA TAAAATAAAA  
 1351 TTTTAGATTT AAGATCACAA GAAATATTAC TGAAGACTTA TACTCTTCC  
 1401 GGGGCTAAGG GAGGTGACAG TCGCTCATCA GAAAAAAA AATGCCCTCA  
 1451 TTTCTTAACCT TTTCTAAAAA ATATAATACA AGTCAGGCT AATACTTCCT  
 1501 GTATATGTGG GAAATTCTA GGGGAAGCTA ACAGGCTTAG AAATAAAGAT  
 1551 GTGTTAAATA GACTACCAAA GTGTCCAATT AAGCAACACG ATACCACCGT  
 1601 TATTGATATT CTAGCAAGAA ATTACTAGCA ATGTTGTAA ATAGACTTAG  
 1651 AAATGCATTT GATGAATTAA CACTTTATA TCTTAATTAA TCTGAATT  
 1701 TCTGTAATGT GAAAATGTT TATTAACTT ATTTCTGGCA TCTATTAGTA

09/9145 49

58/124

1751 AAATTCTGAT GATATACAAG CATTAATATT TTTCCATGGC CACTCAATT  
1801 ATACATACTT TCCCTATCTA TGCTTAGAAG GCAGTGCAA ATTAGATAGT  
1851 AGCAATATTG ATTATAACCA CAAGGTGGAG ACAGATGTCA TGTAAATATGC  
1901 AGTCTGCTCA TATAAAGCAC ATTTCTTAG ACAAGAGTTT TCATACGATA  
1951 TAATAAAGAC ATCTGGAATT TGTCTTGTAT GCAATATGAA ATTTGCTATT  
2001 AACCGTGGAG TTAAAACTTT ATGTCAATAG ATCCAATAAC AATGTTCAT  
2051 AATTAATCAT TATGTCATGC TGTATTTCCA AAAACTATC TTAAATTATA  
2101 AGAGCAAACG AGGTAATAA

Fig. 5 (cont'd)

## Human genomic sequence

1 GTACATTTT TAATAAAGAT GTTTGTGAA ACTTTTGAA TATGAAGATT  
 51 TCTAGTTCTA GAATAATGTT TATAAAAATA TACAAATCCA TCTGGTGATG  
 101 AGTTGACCTC TATCACAACT AGTTGCATA TATAACTTGG GTGTGACCAA  
 151 GCAAGGTGAG AGTTAAGAAC TTTTAAAAC TACTGTATTA TATTGATAGA  
 201 ACTCAGAAAG TACTAACCTG AATATTATTA TTCTAATTGC TTTTCCCTTT  
 251 TAGTTATTAA AAATAAGAAT ACTTAAATTA ATAACAAGAT CTTTTACTGG  
 301 CAGGATTAAC CAAATTATCT GTAATGTGTT CCTCGAATGC TTTTAAGTGG  
 351 AAATATACTT TATACATTCT TTAACAACTC TGAGAGGATG AGTTACATAA  
 401 ATCAGTTCA GAACTATAG AATCTGTAAT ACATAGTAAA GGTTTATTCA  
 451 CAATTAAAAC AATTCACCTT CTATATTAAA AAAACAAATT GTGAAAGTA  
 501 CAGTGGCTTT TCATATGTAT GATTTGTAAA ACAAAATTAGC TTTTTTAAAG  
 551 TGATGTGACG CTTAATGAGA AGAAATCAGT AGAGAATTAC AAACCTGCAC  
 601 TCAAAAGATA CATCTAATAT CATTAAATA ATGAAATTG AAAAAATAGT  
 651 GTGCTCGTT TACAGTCTCA TTAAATGAAT TAAATATCA GCACACATTG  
 701 TAGTAGGTTA TCATTGGCAG AGAAGGCTGA AATAGAAACG TTACAATGGG  
 751 ATGCACTGCC ATCTGAACAT TATGTCGAAG TGGAACGCCGG AAACATATT  
 801 CTCAGAACAA GTGGTAAAT GAAAACAGCA TCATTGTAA AGCATTCTT  
 851 TTGAGAGTGC TTCAGTTCT TCTCCTGATG ACCTGCCATT CAGAAACTGA  
 901 CAATGAATAA TACACTCTGA CACCAGCATT TGTCAATTG CCCAGAACCA  
 951 TATGAGAGTA CTCTAGACAG ATATATGTTC CGAAGTAAAC CGAATACCTG  
 1001 TTAACTGTAA ATCAAATCTT GTAGAAACCA TGCCATGGTT CCTTTGGACA  
 1051 TATACTTGC ATGCCTGAAG CAAGTTACCT TAAGAAATCA TTCTTTGTT  
 1101 TTACAAAAC TGTATTAAAA AATTAAAAAT GCAAAAAGC TTAATATTAT  
 1151 TAGGAATTAA TCCATAGCTT TATTTGGAAT CCAGTTCTT TATTATGATC  
 1201 TATAAACATG CATCATTGA TGGAGTTCT TAGTGGAGAG GTGTTTTCC  
 1251 ATGTTGCTAA GAAACATGCC CCAGCACCAAG AAGGGATACT ACCTACCAC  
 1301 TTTTTGCCAT TTCTCACCGT GATTCTACA TTGTACCTGT TTACTCACTG  
 1351 AACAGGGCTT CCTTCTCTTT GTCTAGATTC TAATCAGGTG TCTTCTGGTG  
 1401 TGGAAGCTT GGCTTTATT TACACACAAAC ACAGAATTAA TAAGATAGAT  
 1451 GCCAAGGATT TAGCAACATT TTAATTCAAC ATTATACAGG TATCAGAGTT  
 1501 AATGAGAATT ATGCATTAGT CTMIAAATT GGGCAGCTTA TTCAGCTAA  
 1551 ACATAGATGT CTAGCTTTA AACACTTTGT TTTTTAATT ACTCTGAAAT  
 1601 TACAATAAAG TCAAAGAACT GAACTGTTT CTTTCAAGC CAGTGAAAT  
 1651 GTGCTTTAGT TATTATTTA CTGGTGATCT AATTATGCAT TTTAATGCTT

09/914549

60/124

1701 TATTACTTAA TACTTATATA AGCCTAAAAT ACGTTGTTAA TGTCTATAATT  
1751 TCAGGGATTT TAGTATTCTT TCCATGAGTT ACCATAACTA GGTGCATATG  
1801 TGTAAATATA CGTATATATC TATATCTATA TATTTATATC TATGTATATA  
1851 TCAATTATA AGACTAAATA GACTTGGCCA TATGTGTTGT TGGTTTATGC  
1901 ATACATGCAC AAATATTGAG GTGTCCACAA AGTATATATG CCTGTACATA  
1951 AATTACATAC TGGCTGGTGA GTGAATGTAAC GCTTCCTCAA ATTGTACAAC  
2001 TCTCCACAGA GTGGCACTCT AATATTGCAA AGGTACAATA TAAGCATGTG  
2051 CAGAATGAAC AGCTCTTCTA GGATCCCTAT AAAACTCCAC CCCATGTTTC  
2101 TGT

Fig. 6 (cont'd)

## Human genomic sequence

1 AAGCTTCATC CCAGAGGGC ACTTGCCAGA TGCCTGCTAG AGCTCTCCTG  
 51 TATGAGGAGT CTATCAACAC CTGCTGGAG GTGTCTCCTC GTCAGGAGGC  
 101 ACGGGGGTCA GGGACCCACT TGAGGAGGCT GTCTGCCCCT TAGCGGAGCT  
 151 AGAACACTGT GCTCGGAGAT CCGCTGCTCT CTTCAGAGCT GGCAGGCAAG  
 201 AGTGTAGGAGT TCTGCTGAGC CTGCGCCAC AGCCGCCCT TCCCCCAGGT  
 251 GCTCTGTCCC AGGGAGATGA GAGTTTATC TGTAAGCCCC TGACTGGGGC  
 301 TGCTACCTTT CTTTCAGATA TGCCCCGCC AGAGAGGAGG AATCTAGAGA  
 351 GGCAGTCTGG CTACAGCAGC TTTGCCAAGC TGCAGTGGC TCTGCCAGT  
 401 CCAAAATTCC CAGCGGGTTT GTTACATTG TGAGGGAAA AGCACCTACT  
 451 CAAGCCTCAG TTATGGCAGT TGCCCCCTCCC CCCACCAAGC TCCAGGGTCC  
 501 CAGGTGTCTC TCAGACTGCT GTGCTGGCAA TGAGAATTTC AAGCCAGTGG  
 551 ATCTTAGCTT GCTGGGCTCC ACAGGGGTGG GATCCACTGA GCTAGACCAC  
 601 TTAGCTCCCT GGCTTCAGCC CCCTTTCCAG GTGAGTGGAT GGTTCTGTCT  
 651 CACTGGCATT CCAGGTGCTA CTGGGGTATG AAAAAAAA CTCCCTGCAGC  
 701 TAGCTTGGTG TCTGCCAGT TTTGTGCTTG AAACTCAGGC CCTTGGTGGT  
 751 GTGGACACCC AATGGAATCT CCTGGTGTGC ATGTTGTGAA GACTGTGGGA  
 801 AAAGCATAGT ATCTGGGCTG GATAGCTCCG TCCTTCAGG CACAGTCCCT  
 851 CATGACTTCC CTTGGCTAGG GGAGGGAGTT CCCCAACCCCT TTGCACTTCC  
 901 CAGGTGAGGC AACACCCAC CCTGCTTCTG CTCACCCCT GTGGGCTGCA  
 951 CCCACTGTCT AATCAGTCAC TGTGAGATGA GCCTGGTACC TCAGTTGGAA  
 1001 ATGCAGAAAT CACCTGCCTT CTGTGTTGAT CTCACTGGGA GCAGCAGACT  
 1051 GGAGCTGTPC CTATTCAGCC ATCTTTCTCA GGTCATAATC ATAGATTTT  
 1101 AATTGATCCC AGCAACATGG ATTGAAAC AGCATATTTC CAAGTGATTT  
 1151 TTTTTTATTT TAAGGTCAAA TCTACAAAAT ATTATAGTGT TATCACCAC  
 1201 TAAAATTATT ACTGGTGATA CTATGTTGT CTCTATTAC ATTTTATTGC  
 1251 TAGAAAAGAT TATAATTGT AGATAATAAT AGTTATTGAA AATGTATTAC  
 1301 ATATCCTTT ACTTTAAGA AGAGGTGACT TAATTATCTA GGTATACAAT  
 1351 TATTTGAGG ATACTAAATG TCATGAATAG CAAATTATC ATATTGCTTT  
 1401 CCTAGGTGAA GACCCTGAAA CAAGAAGAAT GAGAACAGTT AAAAACATAG  
 1451 CAGACTTGAG GCAGAATTAA GAAGAGACTA TGTCCAGTCT TCGTGGGACT  
 1501 CAGATAAGCC ACAGGTTTT TTCAATTG CATATATTG AGCCAATAAA  
 1551 GAAAAAAATAA TTACAAACAA ACATTTAACT TTTCTTATAA TGACAGAGAT  
 1601 GGGATTTCAG TTTCCCCTTA CTATTTCTC CCTTGTGTTA TATCAAATTG  
 1651 ATTGGTAATT ATCCTTAAAC TGAGAATTCA CAGTATATAC CTATTTATCT  
 1701 TTTATCTCTA TCTCTATCTG CTATTTATGT CTTTTTCAGT ATAATTCCA

62/124

1751 GTACTGCAAC TACCACCATC ACTGTTAAGT GGATTGTAA TACCTGTCCT  
 1801 AGAAAACAGT GGCACAAGTT GCACTTGAAA TGCATCTGGG CAGGGTAGTA  
 1851 GGGAGACATT CAAACATAAT TGTAGTTAAC TTTCAGAATA GGTCTGGGAA  
 1901 GGTTACAGTG AGTTAAGGAT TTGTTGAAAA TGTAAACAA TATGTTGTTT  
 1951 TACCCAAAGGT GTACTGATGG CCTTTCTTT GAAAACAAAC GAAAAGCTAT  
 2001 AAAATGTATG CCCCTTTCCA CAATTTGACC TCAAAATGAA TATAGAGTTT  
 2051 AGCTTTCGGG AAGATGACGT GTTTATAAGA GATGACCCCTC AACTCCAGCC  
 2101 TTTTCTGTCT TCATGCATTC TAGATTATGG CCCTAAGTGA ACCAGAGTAT  
 2151 AGTTATTTCT CCATTTTATT TGACAGCACC CTGGAGACAA CATTGACAG  
 2201 CACTGTGACA ACAGAAAGTTA ATGGAAGGAC CATAACCAAC TTGACAAGTC  
 2251 GACCCACCCC CATGACCTGG AGGTTGGGCC AGGCATGTCC GCGACTTCAG  
 2301 GCGGGAGATG CTCCTCCCT GGGTGCTGGC TATCCTCGCA GTGGTACCAAG  
 2351 TCGATTTCATC CACACAGACC CCTCGAGGTT CATGTATACC ACGCCTCTCC  
 2401 GTCGAGCTGC TGTCTCTAGG CTGGGAAACA TGTCACAGAT TGACATGAGT  
 2451 GAGAAAGCAA GCAGTGACCT GGACATGTCT TCTGAGGTG ATGTGGTGG  
 2501 ATATATGAGT GATGGTGATA TCCTTGGAA AAGTCTCAGG ACTGATGACA  
 2551 TCAACAGTGG GTAAGTAACC CTGTTCTCCG TCAGCATTGT GTGAAGAGGG  
 2601 GAGGTGGTCT ACTATAATGC ATTCACTATA AACAAATGTG TAAGTTGCC  
 2651 CAGAAAGTCA TGAGAACATA TGAGATATCT GAGGTTATTG AGAGTGTGA  
 2701 AGGGCCCTTC CTCTGCTCAT TCATGGAGAG TAAAGAATCC AAGATTTCTA  
 2751 TAAATTCTATT ATAAGCCGCT AAGTTTTCT GTTGTGAGA GAAACACATG  
 2801 TGGCTTCTGT TTTTCAGAGT GATTTTCACA TGCTTCTTAA GTAACAGATT  
 2851 TTGTAGTTAA GGACGTGGGA AGGAGACAGG AGGAGTTTG CTGATTTGCT  
 2901 TGATTTTTTT TTTCTTTTT AGCTTGTAG AAGCGGCCTG TAACTGCTTT  
 2951 GAGAAACAAA TATTTCTTA CTGCTTCAA TTATGCATCC CCAATTAAAC  
 3001 TTGAGGGAAA AATCACTTTG GAGTTGAAAG TTTCACTCTA TTCATTTCT  
 3051 TTTGATGGTA TCAGATTCA ATACATCTA GACCCGTGTT TTCTTCTGTG  
 3101 TCCTATTACA TTCCAAAACA TGTTGTGATT GTAAAACCTCT TAGAGTATAT  
 3151 TAACAATTG GGATATTGG CATAATCAGA GAATAGGTCC AAAAGGAGGC  
 3201 AATAGGATAT TCTATTAATA ATTGTAATTG CCATTTTAG CATTCTGTG  
 3251 TATGTACTAT GCTCTTGTCA AGTGTGTTGA AGATAGTGTGTT TTACTTTCC  
 3301 TTCCCACAC CAGCAATGTT TATGAGGTAG ATGTTTTAT ACATGTTCTA  
 3351 TGGATAAGGA AACTGAGTCT AATTGGCCCC GGCTGGGAAC TAACGCTAGG  
 3401 GAAACGGCAG ACCTGCATTA GAACTCAGCT ATGTCTGACT TCAAACACAG  
 3451 GCTCAGTAAT ATGTGGAAAA GCTTCCCAAT TAACTTGTC TATAAACTTT

Fig. 7 (cont'd 1)

099314549

63/124

3501 GTGTGAGTCT GGATTGGAC TTACTCTTG TCTTTACGCA TCTGAGAGGA  
3551 CCCATGTAGG AAATAATTCT TCTATATAAG TGACCCTTCC TGACTTCATT  
3601 CATGAAAAGC TTATGTTGA AGGGTGACAC GACCTAAAAA AGAGTACAAA  
3651 ATAGCTTTG ATTACATTAA TAGCTTGCT CTGATATCCT AATACCTACT  
3701 AGTCCATTCC TGGTATCCAC CCTACCTGAC TTTCTAAAAA TTTAGAATTAA  
3751 TAGAGACTAA TTATGATTAA TTAAGATAGG TTGTTGTTCA GTGCCACTG  
3801 GATTCAGAGT GCCTAGTTG AATCTCTCCC ATTCACTATC TGTGGACCCCC  
3851 TTCGGAACCT AACGTATCCA AATTAGTTTT TGTCACTCTAG AATAAGGATA  
3901 AAATTGTACC ATCTTCATGA AGTTGTTAGG ATCATCCACA AATTTTAGTT  
3951 TGCAGCAATGC TTGGCATGAT ACAAGCACTC AATAAATTAA TCATCTTCCT  
4001 CTTTATCATC ACTATTACAT TTATTATCAT TAATAACCAT ACCAATTCCCC  
4051 GGTTGTTGTT AGTTATAATT ATCATTGGTT TATGTATTAA ACATAGCCTA  
4101 GGAGGCAATG CCCAGTTCAAG AAAACATAAT GGCAGGCAA GAGTGTCTAA  
4151 GGCACACTCT TTCTCCATC TCTCTCTTCT TTCTTCTCCA TTCTTCCAC  
4201 TCTATCCCT CTTCTCTTTT TTTTCTCAAT CTCCTTAGAT GTGGACATAT  
4251 GTGTGAATTG

Fig. 7 (cont'd 2)

64/124

## Human genomic sequence

1 TGTGGGTGTG GGTGTGAAGC ATGTGTATGT GTGTGTGTGA AGCATCTCCC  
 51 CACCTGTAAT GTAAGTCCAT GAGTGCAGAA TTTTGACAT ATTCTTACG  
 101 TGTTGAGTTT TAACAAATGT TTGTGGAGTG AATGAACAAA TTAATGAATA  
 151 TAGGCTATT ATTAAATTAGG CAATATAGTC ACATAGGCTG GCAATCGCAT  
 201 CTAATTAAAT AGAGTGGTAA ATGAGTTCCA GAAAGAACTA AGGTACTACA  
 251 AGGATGTTAT GAAAGAGAAA AATGAGTTAT GTGAAAAATA GGAGACAGTG  
 301 ATAAGAGGGA AAGAATCCC AAGTGTGGC CACATTGTA AACTAATGAC  
 351 CTATTATTCT ATTATTGTTA GCTGAAAGTA GAAAACGTCA TGGGAGGGAA  
 401 TATCTGCTAG TTTTGGTAA AGGATGTTGT GATGGCAGAA CCAAGAAATG  
 451 AACACAAGGT GACTTTGGTT TGGGGACAGT GGGATAATCA ACTCTCCTTG  
 501 CTCCATCAGG GCCCCAGACT GGGCTCTGGC AGAGGAACTC AGAACAAACGT  
 551 AAAGACCTAG ATAGGTATCT AATAAATTGG GACCTGTGAA AACAGTGCCT  
 601 CTTAAAGTGT GGTACCTGGA CCAGCAGCAG CAGCAGCAGC AGCCATTGAA  
 651 ACTTCATAGA AAGACAGATT CTCAGCTTCA TCCAAGACTT ACTGAATTAG  
 701 AATATCTCAA GGTAAGGCCT GGTAATCTGA GCTTTAACTA GCCCTCAAGG  
 751 TGATTCTTAA GTTCAAGCAT CACTATATTA AGTTGAACAA ATAGATGCCA  
 801 GGCCTATAAA TACATGTAAC GCCTAGCATA AATATTCTAA CATTAAAAAT  
 851 GACATTTCAT AGTTCTTATT TACCTATTAA GCTGTGTTCT GTCAAGATAA  
 901 TGAGAATATT GATATGTTAG AATACACTGA TGCACTAATT TTTAAATTAG  
 951 ATCAAATAAT GACTGTTAT ACCTGAAATA AATTGGTTCA GCTTGGTAGA  
 1001 TGCAGTTTT GAGAATTATA TAAGTCATT TTAAAAGAAT AATTTAACT  
 1051 TGAGCTGCTT GCATAAATTAA AATTGCAAAA AGGTCATAGT ATAAATCCTC  
 1101 CTATTAGCAG AGATAGAAGG TTTTAAAAAA AATTACAGAT AAGTCTGAAG  
 1151 GTCTTTAAA ATCTTATATT CAGGAAGTGA CTCGGATGT ATATCATTCT  
 1201 AAAATACATG GTCTTAAATG TTGTAGTTGT ATGACTCTTT CAGTTAATT  
 1251 AAAATACTTC CTTCTATGAA AAATTGTTTC AAAAATTCTT CTAAATTCTG  
 1301 TTATCCATT CAAGTAGGAT AGGCAAGAAC AGATATAAGA TACTACTTT  
 1351 TTGTTCATGT TTACTAAAAA AAAAATTACT GTAATTGAGA TCATGTAAAA  
 1401 ACATGTTCC TGTCTATTG TCTTAACCTT TTAATCCTGG CACCTTAAAT  
 1451 TTGACATAGT AGGAATTAGA AGACAATTGC AGAAAATGTC AACTGGGGAA  
 1501 ATTTTATTCT ACTAAAAACT ATGTCCATAC AACATAGCAA ATCACATT  
 1551 AAAGGCCAAA AAGTCTTCA TAGCAATTTC TCAGATTATT TTCAAAGCAT  
 1601 ATCTTCTCTC TGCTCCTGCA GCATGCCGTT GATTTTCTG TTATGCAGTC  
 1651 ACATAAGTAA TTACATGTTT ACATGTCTAT TTCACTCATA GAACACGAAA  
 1701 CAGTTAAATG TAGAATAATA TCCAATCCAT CTTTTTATCA CCAGTAGCTA

09/914549

65/124

1751 GCATACTGTA GGAACCTCAAT AAATATATCA GATAAATTGT GGAAATAACC  
1801 ATATCAGCTT ATAACATATA GAAATGTGAG TTTAAAAAGA AAACAATTAT  
1851 ACATATGAAA AAATTTTAT ACCATTTTT TAAAGACCTT TCAGATGTCA  
1901 TACAGTTGG ACTTTCCAG TGTTTCTTGT ATCATGAGAC AATAGTAGAC  
1951 ATTGTAAATC AAAAATAGTT TTCTGGGGTT GTGTACATTT GAAAAAACTG  
2001 AATATCATAT CTGTTCTTAG AGAGTAATGA TGGATATTAA CATATCAAAG  
2051 GTACAGAGAA GTCTTAAAGT TCAAAGTAAC ATCTGCTTAA TTGTATTTAA  
2101 TTCAGTGCTC CATGAGCTT TTTATCACTG ATTCCCTCCC TTTTTTCTCT  
2151 TATGATAATA ATTAACTTGT TCCTGTAGCA TTTTAAGAAA TGTTGATTAA  
2201 GTTGAATGCC TTCACTTCTC CAATATAATA GCAGAAACTC AGAAATATTT  
2251 ATTTACCCAG AATCATGCAG CTAATAGTAC AAGGATTCAAG GTCTTTACT  
2301 TCCTATTTG TGGTTCCCAA CTACTTTGC CAAAGGTCTT TTAAATAATA  
2351 TGAAACATAT TAGTGATTGA TTCATTATAG TAAATGGGTA AATGATAAGG  
2401 CTTGCAATAA TTCACTGACA AGAAAGCTT

Fig. 8 (cont'd)

### Murine cDNA sequence

1 AAGCCACAGCACCCCTGGAGACAACCTTGATACGACTGTGACAACTGAAGTGAATGGAA  
S H S T L E T T F D T T V T T E V N G R  
61 GGCCATCCCCAACCTGACAAGCCGACCTCCCCCATGACCTGGAGACTGGGTCAAGCGTG  
A I P N L T S R P S P M T W R L G Q A C  
121 CCCTCGTCTACAGGCTGGAGATGCCCTCCATGGGCCTGGATATTCTCGAAGCGGTAC  
P R L Q A G D A P S M G A G Y S R S G T  
181 CAGCCGATTTCATCCACACGGATCCCTCCAGGTTATGATAACACCACGCCCTCAGGCCGAGC  
S R F I H T D P S R F M Y T T P L R R A  
241 TGCTGTCTCGCGTCTGGAAACATGTCACAAATAGATATGAGCGAGAAAGCAAGCAGTGA  
A V S R L G N M S Q I D M S E K A S S D  
301 CCTGGATGTCTTGAACTGAGCTGGATGTTGGGATACATGAGCGATGGTGAATCCTTGG  
L D V S S E V D V G G Y M S D G D I L G  
361 GAAGAGTCTGAGAGCGGATGATATCAACAGTGGGTACATGACAGATGGTGGGCTCAACCT  
K S L R A D D I N S G Y M T D G G L N L  
421 ATATACCAGAACGCTTAACCGAGTCCCGGACACAGCAACTTCCAGAGATGTCATACAGAG  
Y T R S L N R V P D T A T S R D V I Q R  
481 AGGCCTTCACGATGTGACAGTGGACGCAGACAGCTGGGATGACAGCAGTTCTGTGAGCAG  
G V H D V T V D A D S W D D S S S V S S  
541 TGGCCTCAGTGACACACTTGATAAACATTAGCACAGATGACCTCAACACCACGTCCTCCAT  
G L S D T L D N I S T D D L N T T S S I  
601 CAGTTCTACTCCAACATCACTGTCCTCCAGGAAGAACACTCAGCTGAAAACAGATGC  
S S Y S N I T V P S R K N T Q L K T D A  
661 GGAGAAACGTTGACAGCCATGGTGACGGAGCCAGTGGGATAGTCCTGAGGAGCTGAAGAAAGCCGA  
E K R S T T D E T W D S P E E L K K A E  
721 GGGAGATTGTGACAGCCATGGTGACGGAGCCAGTGGGATAGTCCTGAGGAGCTGAAGAAAGCCGA  
G D C D S H G D G A A K W K G A T S G L  
781 TGCTGAAGACTCGGAGAAGACAGGGCAGAAAGCCAGCCTGCTGTGTCAGACAGGCTC  
A E D S E K T G Q K A S L S V S Q T G S  
841 CTGGAGGAGAGGCATGTCTGCCAGGGAGGAACCTCAGCTACAGCTAGGCAGAAAACAG  
W R R G M S A Q G G T P A T A R Q K T S  
901 CACAAGTGCACCTCAAGACCCCTGGGAAGACAGATGATGCCAAAGCTCCGAGAAAGGGAA  
T S A L K T P G K T D D A K A S E K G K  
961 AACTCCTCTCAAAGGATCATCCTTGCAAAGGTCTCCTTCAGATGCAGGGAAAAGCAGCGG  
T P L K G S S L Q R S P S D A G K S S G  
1021 GGATGAAGGGAAAAAGCCACCGTCAGGCATTGGAGATCGACAGCCAGCAGTTCTTGG  
D E G K K P P S G I G R S T A S S S F G  
1081 ATACAAGAACGCAAGTGGTAGGGCTTCACATGATTACAGCAGCGGTGCCACCAT  
Y K K P S G V G A S T M I T S S G A T I  
1141 CACAAGCGGTTCAGCTACACTGGGAAAATCCCAAATCCGCTGCCATTGGTGGGAAGTC  
T S G S A T L G K I P K S A A I G G K S  
1201 CAATGCAGGAAGGAAAACCAGCTGGACGGTCCCAGAATCAAGATGATGTTGTCCTGCA  
N A G R K T S L D G S Q N Q D D V V L H  
1261 CGTGAGCTCGAAGACCAACCTCCAGTACCGTAGTTGCCCGCCCTCTAAGTCCAGCAC  
V S S K T T L Q Y R S L P R P S K S S T  
1321 CAGCGGAATCCCTGGAGAGGTGGCCACAGGTCGAGCACCAGCAGCATTGATCCAATGT  
S G I P G R G G H R S S T S S I D S N V

Fig. 9

1381 CAGCAGCAAGTCAGCTGGGCCACCACCTCCAAACTGAGAGAACCGACTAAGATCGGCTC  
 S S K S A G A T T S K L R E P T K I G S  
 1441 AGGGCGCTCGAGTCCACTCACTGTCAACCAAACAGACAAAGAGAAGGAGAAAGTAGCAGT  
 G R S S P V T V N Q T D K E K E K V A V  
 1501 GTCAAGATTCAAGAGAGCGTTCTTCAGGTCCCCAAATCCAGCCCCACCTCTGCCAG  
 S D S E S V S L S G S P K S S P T S A S  
 1561 TGCCCTGTGGACTCAAGGGCTCAGACAGCCAGGGTCCAAATATCCAGATATTGCCTCGCC  
 A C G T Q G L R Q P G S K Y P D I A S P  
 1621 CACATTCGAAGGTTTCGGTGCAAGGCAGGGCGCAAATCTGCCCTCCGCACCTAAC  
 T F R R L F G A K A G G K S A S A P N T  
 1681 TGAGGGGGCGAAGTCCTCCTCAGTAGTGCTCAGCCCTAGTACCTCTTAGCCGACAAGG  
 E G A K S S S V L S P S T S L A R Q G  
 1741 CAGTCTGGAGTCACCGTCGTCCGGTACGGGAAGCATGGGAGTGTGCTGGTGGCTGAGTGG  
 S L E S P S S G T G S M G S A G G L S G  
 1801 CAGCAGCAGCCCTCTTCATAAAACCCCTCAGACCTAACATCACAGATTTAAGCTTAAG  
 S S S P L F N K P S D L T T D V I S L S  
 1861 TCACTCCTGGCTTCCAGCCCAGCGTCGGTCACTCTTCACATCCGGTGGCTGTG  
 H S L A S S P A S V H S F T S G G L V W  
 1921 GGCTGCCAATCTGAGCAGTTCCTCTGCCGGCAGCAAGGACACTCCAAGTTACCAAGTCCAT  
 A A N L S S S S A G S K D T P S Y Q S M  
 1981 GACTAGTCCTACGAGCTCTGAGTCATTGACCTGCCCTCAGCCATCATGGCTCCCT  
 T S L H T S S E S I D L P L S H H G S L  
 2041 GTCTGGACTGACCACAGGGCACTCACGAGGTGCAGAGCCTGCTCATGAGAACGGTAGTGT  
 S G L T T G T H E V Q S L L M R T G S V  
 2101 GAGATCTACTCTCTCAGAAAGATACACCCCCATCATCTGGCAGGCCAACCAAGAACAG  
 R S T L S E R Y T P S S R Q A N Q E E G  
 2161 CAAAGAGTGGCTGCGATCGCATTCCACTGGGGCTGCAGGATACTGGCAACCAGTCTCC  
 K E W L R S H S T G G L Q D T G N Q S P  
 2221 CTTGGTCTCCCTCTGCATGTCATCGTCAGCCACCGGAAAATATCACCTTCAACTT  
 L V S P S A M S S S A T G K Y H F S N L  
 2281 GGTGAGTCCCACCAACCTCTCCAGTTAACCTGCCTGCACCCAGTATGATGCGCTCCAG  
 V S P T N L S Q F N L P A P S M M R S S  
 2341 CAGTATCCCCGCCAGGACTCCTCCCTCGACCTCTATGATGATGCCAGCTTGGTAG  
 S I P A Q D S S F D L Y D D A Q L C G S  
 2401 TGCAACTTCCCTGGAGGAAAGGCCACGGGGCTTAGCCACTCCGGCTCATTCAAGAGACAG  
 A T S L E E R P R A V S H S G S F R D S  
 2461 CATGGAGGAAGTTCATGGCTTCACTGTCATTGGTCTCCAGCACATCATCCCTTACTC  
 M E E V H G S S L S L V S S T S S L Y S  
 2521 TACGGCTGAAGAGAACGGCTATTAGCAAGAACATCCATAAGCTACGGAGAGAACTGGTTGC  
 T A E E K A H S E Q I H K L R R E L V A  
 2581 CTCCCAGGAGAAAGTCGCTACCCCTACGTCTCAGCTGTCAGCAAATGCTCACCTGTAGC  
 S Q E K V A T L T S Q L S A N A H L V A  
 2641 AGCTTTGAAAAGAGTTAGGGAAATATGACTGGCCGTTGCAAAGTCTAACCATGACAGC  
 A F E K S L G N M T G R L Q S L T M T A  
 2701 GGAACAAAAGGAATCTGAGCTTATCGAACTGCGGGAAACCATGAAATGTTGAAGGCCA  
 E ..Q K E S E L I E L R E T I E M L K A Q

Fig. 9 (cont'd 1)

090314549 060502  
09/914549

68/124

2761 GAACTCTGCTGCCAAGCAGCCATTCAAGGGAGCACTGAATGGCCAGACCACCCCTCCAA  
N S A A Q A A I Q G A L N G P D H P P K  
2821 AGATCTCCGCATCAGAAGACAGCACTCCTCTGAAAGTGTTCAGTATCAACAGCGAAC  
D L R I R R Q H S S E S V S S I N S A T  
2881 GAGCCATTCCAGCATGGCAGTGGTAATGATGCTGACTCCAAGAAA  
S H S S I G S G N D A D S K K

Fig. 9 (cont'd 2)

## Murine genomic sequence

1 GGGATGAAGG GAAAAAGCCA CCGTCAGGCA TTGGAAGATC GACAGCCAGC  
 51 AGTTCTTTG GATAACAAGAA GCCAAGTGGT GTAGGGGCTT CCACTATGAT  
 101 TACCAGCAGC GGTGCCACCA TCACAAGCGG TTCAGCTACA CTGGGGAAAA  
 151 TCCCCAAATC CGCTGCCATT GGTGGGAAGT CCAATGCAGG AAGGAAAACC  
 201 AGCCTGGACG GGTCCCAGAA TCAAGATGAT GTTGTCTGC ACGTGAGCTC  
 251 GAAGACCACC CTCCAGTACC GTAGTTGCC CCGCCCTTCT AAGTCCAGCA  
 301 CCAGCGGAAT CCCTGGGAGA GGTGGCCACA GGTGAGCAC CAGCAGCATT  
 351 GATTCCAATG TCAGCAGCAA GTCAGCTGGG GCCACCACCT CCAAACGTGAG  
 401 AGAACCGACT AAGATCGGCT CAGGGCGCTC GAGTCCAGTC ACTGTCAACC  
 451 AAACAGACAA AGAGAAGGAG AAAGTAGCCAG TGTCAGATTG AGAGAGCGTT  
 501 TCCTTGTCAAG GTTCCCCCAA ATCCAGCCCC ACCTCTGCCA GTGCCTGTGG  
 551 GACTCAAGGG CTCAGACAGC CAGGGTCCAA ATATCCAGAT ATTGCCTCGC  
 601 CCACATTTCG AAGGTAAGGG TATGAAAGA GATGTTGGGAA AACATAAAA  
 651 GGTAGTATAT AGCATGTATT TATTCTGTAC GAAACTATTT TCATGTATTG  
 701 TAAATATTCT AAGATTCTGT ATCTTATACT TGTCTAAAAT ATAGTGATTT  
 751 TATTTTGCTG ATTGCACCTG TTGCTAGTGT AAAAGCATTG CTCATTTAGA  
 801 GAGTGGTTAG CCTTCAGCT ATACAGCCAG TGTGACACTA AAATACAGAT  
 851 ACCACTTGTA GCGGGCATAA AACCAACATGA CTGACTATTG ATAGAAATAA  
 901 AGTGATAGCT TGAAAGATA TTTAGTGATT TCCACCTCTC CTTTCCAGAA  
 951 TTAAAAAAAG CAAATTGCAT AGATCTTAT AAACACATTG ACTTCTAGTG  
 1001 TATGTTATCT TGTTGACTCT TAATGAAATG GCAGTTATGA ATATAGATGA  
 1051 TATATTCTTT CTAACAGTTT ATAAGAGACC AATTATACAGA GTACCAGATC  
 1101 TTAACATAGT ACAATAACA GCAACAAAAA CAACCCAAAA AGCTATCAA  
 1151 GTATGGCTG ATTGCAGAAT TTGAAAACAT TTACATGTTT GACATAGGAC  
 1201 AAGAACTCAG GAGTGAGGTG ACTTTTTATA AGTCTTCATC AATGTCCTT  
 1251 TACAGGAACC AGGAAGCATA TCTGATATAT GTGTCAGGAT TATCACTTTA  
 1301 TTAATTATGT GAAATTCTGT TTAGAAATCT ACCTGATTAA AAATACTTA  
 1351 ATATAGTAGG GGTAAAATT AGTTAATGAG TTAAGACAAG TTGTTAAATA  
 1401 ATCCTGGCTC TGTTTTCTCA TCTTCAAAT GATAGAGTAT AATTATCAC  
 1451 CTCTTGTAA ATATTCAGG TTTGTGTTA TTCTCTTGAT AACTTTGATC  
 1501 TCTTGAAGA GTCTGAAGA ATTTACATTA AGTAATCTTA GAAACATAAC  
 1551 TATTTGAGAA ACAGTAGTCA AATTTGTCA TTAGAAGTAT TAACTCTGAA  
 1601 GAATGATTG AAGTGACAGT TCTTAGAAAG AATAAATTAT AGCTTGTAGC  
 1651 AAGAGTAAAT ATTTCACTG CTTGTGTGAG AGCCAAGAGC GCCCTCTTGT  
 1701 GGCCCATTAC CTATGAAACA ATTTCTCATA TTGCCCCAGA AAATCTTCCA

09/914549

70/124

1751 CTGCAGGAAA TAATGGATTT CATTGCCTCT GAATTAGTAA CCATTCTGCC  
1801 ATTTCTTCAT ACCATTTTAT TTCCATACTT GCATAAAATTT GATTATGTCA  
1851 TCTGCTTCAT TTACAAAAACT AAAATGTTT CTGAGCTAAA CTCCAGTAGC  
1901 TAACTTAGTA CAAATGGTAT TTTAAATCA CTGCTATAAG TATATATATT  
1951 TGAATAGCTC TGGCAACGGA CGGAAATCCC TATGGTCTTT CCATGGGAAG  
2001 ATACAAACCA ATCCATAAGT TGTCCAGCAA TATCCAATAT TTCCAGCCCCA  
2051 GCCAGTCAGG CCTCTTAAAC ATTACCTTAC ATATTGAAAC CTTTCCTTAA  
2101 ATGTCCCCTT TAGACAATCT ATTTTTTAAA AAGATGAAAA TCCATTAAAG  
2151 CATCATATAT CGAATGCGTA GAAGTTGTTT CATTATAATG GTTCTGCAGA  
2201 TAGGTAATGC CAAAACGGCC AAAATATTG ATCACTAGAA GCGTAAAAGT  
2251 CAAGTACAAT CATGTTGACT TTTTTTCCAA GGTGGGTTCA CTGCTGCCCA  
2301 CCTTGTTCC AGGCCAGTGC TTACTTAAGA TATCGTAAGT GATTTTTTTT  
2351 TAATTTTAA TTTTTTAGTA GTTGGTTAAT CAAAAGCCAG TCATGTCACC  
2401 TTCAGGAACA TAGAGGCTGG ACGTGCTTGG CAGCTCACGA CTCCAAAGCA  
2451 CACTTGGCTC TGTGGACTGA AACCCCTAGGA AACGTGGATG TGAGTCTCTT  
2501 GGAACAACTC AAGTTGTTAT TTGTTTTCT TTTAGGTTGT TCGGTGCCAA  
2551 GGCAAGGGGGC AAAATCTGCCT CCGCACCTAA TAC

Fig. 10 (cont'd)

T2HC

Homologous human cDNA

1 GGATCAGCTTCGGGAGACCATGCACAACATGCAGTTGGAGGTGGACCTGCTGAAAGCAGA  
     D Q L R E T M H N M Q L E V D L L K A E  
 61 GAATGACCGACTGAAGGTAGCCCCAGGCCCTCATCAGGCTCCACTCCAGGGCAGGTCCC  
     N D R L K V A P G P S S G S T P G Q V P  
 121 TGGATCATCTGCATTATCTTCCCCACGCCGCTCCCTAGGCCTGGCACTCACCCATTCCCT  
     G S S A L S S P R R S L G L A L T H S F  
 181 CGGCCCCAGTCTTGAGACACAGACCTGTCACCCATGGATGGCATCAGTACTTGTGGTCC  
     G P S L A D T D L S P M D G I S T C G P  
 241 AAAGGAGGAAGTGACCCCTCCGGGTGGTGGAGATGCCCGCAGCACATCATCAAAGG  
     K E E V T L R V V V R M P P Q H I I K G  
 301 GGACTTGAAGCAGCAGGAATTCTCCTGGGCTGTAGCAAGGTCACTGGAAAAGTTGACTG  
     D L K Q Q E F F L G C S K V S G K V D W  
 361 GAAGATGCTGGATGAAGCTGTTCCAAGTGTCAAGGACTATATTTCTAAATGGACCC  
     K M L D E A V F Q V F K D Y I S K M D P  
 421 AGCCTCTACCCCTGGGACTAACGCACTGAGTCCATCCATGGCTACAGCATCAGCCACGTGAA  
     A S T L G L S T E S I H G Y S I S H V K  
 481 ACGAGTGTGGATGCCAGAGCCCCCGAGATGCCCTGGCGTCGAGGTGTCAATAACAT  
     R V L D A E P P E M P P C R R G V N N I  
 541 ATCAGTCTCCCTCAAAGGTCTGAAGGAGAAATGCCGACAGCCTGGTGTGAGACGCT  
     S V S L K G L K E K C V D S L V F E T L  
 601 GATCCCCAAGCCGATGATGCCAGCACTACATAAGCCTCCTGCTGAAGCACCAGCGCCTCGT  
     I P K P M M Q H Y I S L L L K H R R L V  
 661 CCTCTCGGGCCCCAGCGGCACGGCAAGACCTACCTGACCAATCGCTTGGCCGAGTACCT  
     L S G P S G T G K T Y L T N R L A E Y L  
 721 GGTGGAGCGCTCTGGCCGTGAGGTACAGAGGGCATCGTCAGCACCTCAACATGCCA  
     V E R S G R E V T E G I V S T F N M H Q  
 781 GCAGTCTTGCAGGGATCTGCAACTGTATCTTCAACCTAGCCAACCAGATAGACCGGGA  
     Q S C K D L Q L Y L S N L A N Q I D R E  
 841 AACAGGAATTGGGATGTGCCCTGGTATTCTATTGGATGACCTGAGTGAAGCAGGCTC  
     T G I G D V P L V I L L D D L S E A G S  
 901 CATCAGTGAGTTGGTCAATGGGCCCTCACCTGCAAGTATCATAAATGTCCCTATATTAT  
     I S E L V N G A L T C K Y H K C P Y I I  
 961 AGGTACCAATCAGCCTGTAAAAATGACACCCAACCAGGGCTTGACCTGAGCTCAG  
     G T T N Q P V K M T P N H G L H L S F R  
 1021 GATGTTGACCTTCTCCAACAAACGTGGAGGCCAGCAATGGCTCCTGGTGTACCTGAG  
     M L T F S N N V E P A N G F L V R Y L R  
 1081 GAGGAAGCTGGTAGAGTCAGACAGCGACATCAATGCCAACAGGAAGAGCTGCTTGGGT  
     R K L V E S D S D I N A N K E E L L R V  
 1141 GCTCGACTGGTACCCAAGCTGTGGTATCATCTCCACACCTTCTTGAGAAGCACAGCAC  
     L D W V P K L W Y H L H T F L E K H S T

09/914549

72/124

1201 CTCAGACTTCCTCATCGGCCCTGCTCTTCTGCGTCCCATTGGCATTGAGGACTT  
S D F L I G P C F F L S C P I G I E D F  
1261 CGGGACCTGGTTCATTGACCTGTGGAACAACCTATCATTCCTATCTACAGGAAGGAGC  
R T W F I D L W N N S I I P Y L Q E G A  
1321 CAAGGATGGGATAAAGGTCCATGGACAGAAAGCTGCTGGGAGGACCCAGTGGATGGT  
K D G I K V H G Q K A A W E D P V E W V  
1381 CCGGGACACACTTCCCTGGCCATCAGCCAAACAAGACCAATCAAAGCTGTACCACCTGCC  
R D T L P W P S A Q Q D Q S K L Y H L P  
1441 CCCACCCACCGTGGGCCCTCACAGCATTGCCTCACCTCCGAGGATAGGACAGTCAGAACAGA  
P P T V G P H S I A S P P E D R T V K D  
1501 CAGCACCCCCAAGTTCTCTGGACTCAGATCCTCTGATGGCCATGCTGCTGAAACCTCAAGA  
S T P S S L D S D P L M A M L L K L Q E  
1561 AGCTGCCAACTACATTGAGTCTCCAGATCGAGAAACCATCCTGGACCCCAACCTCAGGC  
A A N Y I E S P D R E T I L D P N L Q A  
1621 AACACTTTAAGGGTTCGGCAATCACTGTACCCCCGGACAGCAGAACGCTGGCATCAGCT  
T L \*  
1681 ATCTTAGCTCCTCCTCTCCCTCTCCTCTTCAGAGCACTGGCTCTCCAGCCCCAGGAGG  
1741 AGAACAGGAGGGAGGAGAGATGAAAGAGGGAGGGACAGGTTCTGGTGTACCTTGA  
1801 GAACTTCCTAGGAAGGAATGGTGGGTGGCGTTGGAACTTGTGCCCCCTAAACACATT  
1861 TACTGGCCTCCTCTAAATGACTTGGGAAAAGATGATTCTGGTCTTCCCTGACTTCT  
1921 TGTTCAATTACAAACTCCTGGCTTCTGGGAGGGTTCAAGAAAACATCAAAACACTG  
1981 CAGCAGTTCTAAATGATTCTCACAAGCAACCCCTGAGAGAGACAGTCTTGTGAGGGAGAT  
2041 CTGGGGAGGCAGGAAGCTCCTCAGATTTCTCACAGACCCCTCCAATTCCATCACCAC  
2101 TGCCAACAACTCCTCCCCAGAGATCTGGCTGGAGCCCAGAAAAAGAACATGTGGTTA  
2161 AAAATGTTAAATCAATCTGAAAAGTAAAATGAAAAACAAAACAAGCAAACAAAC  
2221 AAAAACAAATGGAAAAGATGAAGCTGGAGAGAGAGGAACAGTTGCCAAGGTAGAGAGCT  
2281 GCCCGCTCCTGCCCTCTGGATGACATAGGGACATCAACAAGACGGCTGCCAACCTGAGA  
2341 AGTCACCAAACCACAAAATAACCTTACAGCCTTCAGGGAAAGACTACCAGCTGTCTT  
2401 TCTACCCTCTAAATTAAACAATGCATAAGAGTCATAAAACCTACTTTTAAAAAA

Fig. 11 (cont'd)

09/914549

73 / 124

Homologous murine cDNA sequence

1 GAACTATGGGAAAAAGAGATGAAGCTCACGGATATCCGGTTGGAGGCCCTCAACTCTGCC  
E L W E K E M K L T D I R L E A L N S A  
61 CACCAGCTGGACCAGCTCGGGAGACCATGCACAATATGCAGTTGGAGGTGGACCTGCTG  
H Q L D Q L R E T M H N M Q L E V D L L  
121 AAACAGAGAATGACCGGCTGAAGGTTGCCCGGCCCTCAGGCTGCACTCCAGGG  
K A E N D R L K V A P G P S S G C T P G  
181 CAGGTCCCTGGGTCACTCGGCTCTGTCGTCGGCTGCAGCTCCCTGGGCTTGCACTCAGC  
Q V P G S S A L S S P R R S L G L A L S  
241 CATCCTTCAGTCCTAGTCTCACAGACACAGACCTCTCACCCATGGATGGCATCAGCACC  
H P F S P S L T D T D L S P M D G I S T  
301 TGTGGTTCAAAGGAAGAGGTGACCCCTGGGGTGGTCCGGATGCCGCCCCAGCACATC  
C G S K E E V T L R V V V R M P P Q H I  
361 ATCAAAGGGACTTAAAGCAGCAGGAGTTCTTCCCTGGGTCAGCAAGGTCAAGTCAGGGCAA  
I K G D L K Q Q E F F L G C S K V S G K  
421 GTGACTTGGAAAGATGCTGGATGAAGCCGTTTCCAAGTGTCAAGGACTACATTCTAAA  
V D W K M L D E A V F Q V F K D Y I S K  
481 ATGGACCCAGCCTCAACCTGGGACTGAGCACTGAGTCCATACATGGCTATAGCCTCAGC  
M D P A S T L G L S T E S I H G Y S L S  
541 CACGTGAAACGAGTGCTGGATGCTGAGCCCCAGAGATGCCCTCTGCCGCCAGGGTGC  
H V K R V L D A E P P E M P P C R R G V  
601 AATAACATATCAGTCGCTCTCAAAGGTCTGAAAGAGAAGTGTGTCGACAGCCTGGTT  
N N I S V A L K G L K E K C V D S L V F  
661 GAGACGCTTATCCCCAACGCCATGATGCAGCACTACATCAGCCTCTGCTCAAGCACCGG  
E T L I P K P M M Q H Y I S L L L K H R  
721 CGCCTGGTGCCTCCGGCCCCAGTGGCACCGGCAAGACCTACTTGACCAATCGGCTAGCC  
R L V L S G P S G T G K T Y L T N R L A  
781 GAGTACCTGGTGGAGCGCTCCGGCGAGGTACGGATGGCATCGTCAGCACTTCAAC  
E Y L V E R S G R E V T D G I V S T F N  
841 ATGCACCAGCAGTCTGCAAGGATCTGCAACTGTACCTCTCCAACCTAGCCAACAGATA  
M H Q Q S C K D L Q L Y L S N L A N Q I  
901 GACCGGGAAACAGGGATAGGGATGTGCCCTGGTATCCTGGATGATCTGAGTGAA  
D R E T G I G D V P L V I L L D D L S E  
961 GCAGGCTCCATCACTGGTCAATGGGCCCTCACCTGCAAGTATCACAAATGTCCC  
A G S I S E L V N G A L T C K Y H K C P  
1021 TACATTATAGGTACCAATCAGCCTGTAAAAATGACACCCAACCATGGCTTGCACCTTG  
Y I I G T T N Q P V K M T P N H G L H L  
1081 AGCTTCAGGATGCTGACCTCTCGAACATGTGGAACAGCCAATGGCTTCTGGTCCGT  
S F R M L T F S N N V E P A N G F L V R  
1141 TACCTGGAGGAAGTGGTAGAGTCAGACAGTGCAGTCAATGCTAACAGGAAGAGCTG  
Y L R R K L V E S D S D V N A N K E E L  
1201 CTTCGGGTGCCTGGACTGGGTGCCCAAGCTGGTATCACCTCCACACCTTCCGGAGAAG  
L R V L D W V P K L W Y H L H T F L E K  
1261 CACAGCACCTCGGACTTCCCTATTGGCCCTTGCTTCTCCCTGTCCCTGCCCCATTGGCATC  
H S T S D F L I G P C F F L S C P I G I  
1321 GAGGACTTCCGGACCTGGTCATTGACCTGTGGAACAATTCCATCATCCCCTATCTACAG  
E D F R T W F I D L W N N S I I P Y L Q

74 / 124

1381 GAAGGAGCCAAGGATGGGATCAAGGTTCATGGACAGAAAGCTGCTTGGGAAGACCCGGTG  
 E G A K D G I K V H G Q K A A W E D P V  
 1441 GAATGGGTCCGAGACACTCTTCCCTGGCCGTCGGCCCAACAAGACCAATCAAAGCTCTAC  
 E W V R D T L P W P S A Q Q D Q S K L Y  
 1501 CACCTGCCCGCCTCTGTGGGCCCCACAGCACTGCCTCACCCCCGGAGGACAGGACA  
 H L P P P S V G P H S T A S P P E D R T  
 1561 GTCAAAGACAGCACTCCAAACTCCCTCGACTCAGATCCCTGATGGCCATGCTACTGAAA  
 V K D S T P N S L D S D P L M A M L L K  
 1621 CTCCAAGAAGCTGCCAACTACATTGAGTCACCAAGATCGAGAGACTATCCTGGACCCCAAC  
 L Q E A A N Y I E S P D R E T I L D P N  
 1681 CTCCAGGCAGCACTCTGAGGGCCCGCAGTCACTGTCACCCCTGGAGGGCAGAAGGCTGGC  
 L Q A T L \*  
 1741 TTCAGCATCATTAGCTCCTCTGCCCTTCCCTCATAGCTCTGGCTCACAGCCTCGC  
 1801 CAAGAGAACAGGAGGGAAAGAAGAGGGCAGGAGGGATGGTTCTCGGTGCTGAACCTT  
 1861 TGAGAACTTCCTACTAGGAATTGGAGGGGTGGAGTTGAGAACTCCGTGCCCTTAACCT  
 1921 ACATTTGCTGGCCTCCTCTTACGACTTAGGAGAAAAGATGATTCTGGCTTTCTTCAG  
 1981 TTTTGTTCACCTACAAACTCTTGGCTTCTGGGAGGGATCGGAAGATATAAACAGA  
 2041 CAAACAAAACAAACAAACCAACTACAGCAGTTCAAGCTCGTTCTACAAACACCTCTG  
 2101 AGACAGTCACATGTGGCAAATCTAAGGGAGGCAGGAAGCTCTACAGACTTTCTGCAA  
 2161 CCCTTCCCAGTTCTGTCGACACTGCCAACACCTCCCCGCCAGAGACCTGGCCAGGCCA  
 2221 AGAAAAGAGAACATGTGGTTAACAGAAAAACAAAACAAAACAAAATATATG  
 2281 TGTAAATCAACCTGTAGAAGGTAAAACGGCAATGGAAAAGATGAAGCTGGAAGGAGGG  
 2341 CCCAGTTGCCAAGATGGAACGAGAGCTGCCAGATCTGCCTCTGGATGACAAGAGGG  
 2401 CATTGCAAGATGGCTGCCAGTCTAAACGTCAACAGACCACAAAGAGTAACATCACAGCCT  
 2461 TCGAAGAAAGGCCACAAGCTGTCTTCTGCCCTCTAACATGAACATGCATGAAAAGTCAAT  
 2521 AAACCTACTTTTAATTTTAAAAAAAAAAAAAAAAAAAAA

Fig. 12 (cont'd)

09914549

75/124

T2 Murine cDNA with following intron

CCAATAGAACTCCGGATCAAGAGGCAGAATTCCCTCAGATAGCATCTCCAGCCTAACAGC  
 1 P I E L R I K R Q N S S D S I S S L N S -  
  
 ATCACCGCCATTCCAGCATCGGCAGCAGCAAAGATGCTGATGCCAAGAAGAAAAAGAAG  
 61 -----+-----+-----+-----+-----+-----+-----+-----+ 120  
  
 I T S H S S I G S S K D A D A K K K K K -  
  
 AAGAGTTGGGTATGTAAAGGCTTGGGATCGGCCTGTGCTAGGAGTCACTCACCTGTTG  
 121 -----+-----+-----+-----+-----+-----+-----+-----+ 180  
 K S W  
  
 CAGGGAACTGACCCCTTCAGGATCAACAAAGAGGGTCCCTCTAACAGGATGCCAGTGT  
 181 -----+-----+-----+-----+-----+-----+-----+-----+ 240  
  
 TGTGACATCTGCTGGGGACAAAATTCACTAAGTCCCATTCTCTATCCATTGTCTATT  
 241 -----+-----+-----+-----+-----+-----+-----+-----+ 300  
  
 CTCCTTACCAACGCCCTGCACATATAACCCAGCCCCCACCCTGCATCCTTATAC  
 301 -----+-----+-----+-----+-----+-----+-----+-----+ 360  
  
 ATGTCTGCTATCCTGGGCTCTACCTACTGATGAGGTCAAATGTATTGGCCGTAGAAGG  
 361 -----+-----+-----+-----+-----+-----+-----+-----+ 420  
  
 AGCTGAGAAAATTATTCATGGTGGAGAGTGGGCATGTGGAGAGAATTGTAAGCCAA  
 421 -----+-----+-----+-----+-----+-----+-----+-----+ 480  
  
 GCAGGGTACTCTAGACGCTCCTGGGCTGTTGCTTAGTTGGGTGAGGAGGCTGTGGAA  
 481 -----+-----+-----+-----+-----+-----+-----+-----+ 540  
  
 CGTCCCCATCGCTCCAAGCCTGCTTTGTCTGGTCCAGAGGTGGTTGTTCTGTGTGG  
 541 -----+-----+-----+-----+-----+-----+-----+-----+ 600  
  
 TATCCCCCTGTAACTCTAAACTGGCTTGTTGGTGAGCTTCTACAATCTGTACGCAGGTG  
 601 -----+-----+-----+-----+-----+-----+-----+-----+ 660  
  
 TAGGGCACTGCCTGACTGACTGAAAGGGAGAGTGACCCAGAGTGAGCGTCTGTCCCTGT  
 661 -----+-----+-----+-----+-----+-----+-----+-----+ 720  
  
 CCCTGCTGAGGAGGGCTGGCTACAGACTTGGCTAGTGCAGACAGGAGCCAGCTGTGTG  
 721 -----+-----+-----+-----+-----+-----+-----+-----+ 780  
  
 GAGAAGCAGCTGTGAAATGCATGAGTAGTGTGCTGCTGCTGCTGCTGCTTCTT  
 781 -----+-----+-----+-----+-----+-----+-----+-----+ 840  
  
 TTCATTGTTTTTTTTTTCTTCTTCTTCAAAATGCTGACCTCAAATC  
  
 CCTATTTTTTCCAGGTTATGAGGTAAGAAACTCGGTTCTCCTCGTGTGCTTTCT  
 901 -----+-----+-----+-----+-----+-----+-----+-----+ 960  
  
 TTCCCTTGCACACCTCGTGTCCAGAGCAAGCACCTCTCTTAAAAAAAAAAAAAAA  
 961 -----+-----+-----+-----+-----+-----+-----+-----+ 1020  
  
 AAAAA  
 1021 ----- 1025

Fig. 13

09/914549

76/124

splicing variant 1 (JFC410)

1 AGCGAGTTACTCACGCTTCCCATCGGAAGCCAGCCAGGCCAAACCCAGCAAGATA  
R V T H A S P P S E A S Q A K T Q Q D M  
61 TGCAGTCAGTCTGGCAGCCAGATATGCAACTCAGTCTAATCACAGTGGATTGCAACCA  
Q S S L A A R Y A T O S N H S G I A T S  
121 GTCAAAAAAAAGCCTACTAGGCTTCCAGGGCCCTCTAGGGTGCCTGCTGCAGGAAGCAGCA  
O K K P T R L P G P S R V P A A G S S S  
181 GCAACCTCCACCCACCCCTTATTTAAATAGGAGAAGTCAGAGCTTAACAGCATTGACA  
K V Q G A S N L N R R S Q S F N S I D K  
241 AA

bp 1 corresponds to bp 914 of THC

underlined sequence represents further splicing form and is not shown in the THC sequence

Fig. 14

09/914549

77/124

splicing variant 2

1       GGCACTCACGAGGTCCAGAGCCTGCTCATGAGAACGGTAGTGTGAGATCTACTCTCTCA  
      G T H E V Q S L L M R T G S V R S T L S  
  
61       GAAAGATATAACCCCATCATCTCGGCAGGCCAACCAAGAAGAGGGCAAAGAGTGTTGCGT  
      E R Y T P S S R Q A N Q E E G K E W L R  
  
121      TCTCATTCTACTGGAGGGCTTCAGGACACTGGCAACCAG  
      S H S T G G L Q D T G N Q

bp 1 corresponds to bp 3300 of THC

underlined base pairs -> position of the differentially spliced exon which lacks here but is shown in the THC sequence

Fig. 15

## T2-cDNA sequence and T2 protein encoded therein

CCGCGGGGCTTCCATCCTCCTTGACTGATTTAAATTAAATTGTATTTCCCCGC  
 1 R G A S I L P L T D F \* I L I C I F P A - 60  
  
 CGCCCCGCCCTTTCCCTCCGACCCCGCCCTATCGCTCCCCGGCTCCCTGCTCTTCCT  
 61 A P P L F L R P R P I A P R L P C S F L - 120  
  
 TTTTCCCGGCTTCCTCCTCGCGTTCTTCCCCTCGGCCCTCGGCTTGCCCTCTCCCT  
 121 F P G F L P R V S F P C A L G L P L S L - 180  
  
 CCTCCCTCGCTCTCTCCCCCTCTCTCCCCCTCTCCCTCGGTTCTCCGTCCCTCTCT  
 181 L P R S L P L S P S S S V S S V L S L - 240  
  
 CCCCCCTCCTCCCTCCCCCGCCTCCTCCTCGCTCCGCCCTGCCCTCCCCCGT  
 241 P L L L P R L L L R S R P L P P P P V - 300  
  
 GCCTGCAGACGCGCGGATCGTCCATGCGCTCCTCGCGGGCAGAATGCTGGCAGCAGCGT  
 301 P A D A R I V H A L L A G R M L G S S V - 360  
  
 CAAGAGCGTGCAGCCCGAGGTGGAGCTGAGCAGCGCGGCGACGAGGGCGCGACGA  
 361 K S V Q P E V E L S S G G G D E G A D E - 420  
  
 ACCCGGGGGCGCCGGCAGGAAGGCGGCAGCGGGACGGCAGAGGCATGCTGCCAAGCG  
 421 P R G A G R K A A A A D G R G M L P K R - 480  
  
 CGCCAAGGCGCCGGCGGCCGGCATGGCAAGGCCAGCGCGCTGAGCTGAAGGT  
 481 A K A P G G G G M A K A S A A E L K V - 540  
  
 CTTCAAGTCCGGCAGCGTGGACAGCCGTGTCCCCGGCGGCCGCGCTCCAACCTGCG  
 541 F K S G S V D S R V P G G P P A S N L R - 600  
  
 CAAGCAGAAGTCACTCACCAACCTCTCTTCTCACGGACTCCGAGAAAAAGCTGCAGCT  
 601 K Q K S L T N L S F L T D S E K K L Q L - 660  
  
 TTATGAGCCCAGATGGAGCGACGATATGGCAAGGCCAAAGGCTTAGGCAAGGTGGG  
 661 Y E P E W S D D M A K A P K G L G K V G - 720  
  
 GTCCAAGGGCCGTGAAGCTCCGCTGATGTCCAAGACGCTGTCCAAGTCGGAGCACTCGCT  
 721 S K G R E A P L M S K T L S K S E H S L - 780  
  
 CTTCCAGGCCAACGGGAGCCCGGGCGGCCAACAGACCCCCCTGGCTCCGCTCGCGCC  
 781 F Q A K G S P A G G A K T P L A P L A P - 840

097914549

79/124

CAACCTGGGAAAGCCGAGCCGGATCCCTCGAGGACCCATGCGGAGGTCAAGCCGCTCAG  
841 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 900  
N L G K P S R I P R G P Y A E V K P E S -  
  
CAAGGCGCCTGAAGCGGCCGTGAGCGAAGATGGCAAATCGGACGACGAGCTGCTCTCCAG  
901 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 960  
K A P E A A V S E D G K S D D E L L S S -  
  
CAAGGCCAAGGCAGCAAAAGAGCTCTGGCCTGTCCCCTGCCAAGGGCAGGAGGAGCG  
961 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1020  
K A K A Q K S S G P V P S A K G Q E E R -  
  
CGCCTTCCTCAAGGTGGACCCGAGCTGGTGGTGACCGTGCTGGAGACCTGGAGCAGCT  
1021 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1080  
A F L K V D P E L V V T V L G D L E Q L -  
  
GCTCTTCAGCCAGATGCTGGACCCAGAGTCCCAGAGAAAGAGGGACAGTGCAGAATGTCCT  
1081 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1140  
L F S Q M L D P E S Q R K R T V Q N V L -  
  
GGATCTCCGGCAGAACCTGGAAGAGACCATGTCCAGCCTGCGAGGGTCCCAGGTGACTCA  
1141 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1200  
D L R Q N L E E T M S S L R G S Q V T H -  
  
CAGCTCCCTGGAGATGACCTGCTACGACAGCGATGCCAACCCACGCAGCGTGTCCAG  
1201 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1260  
S S L E M T C Y D S D D A N P R S V S S -  
  
CCTCTCCAACCGCTCGTACCCCTGTATGGCGCTATGCCAGTCCAGTCCGGCTGCA  
1261 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1320  
L S N R S Y P L S W R Y G Q S S P R L Q -  
  
GGCTGGTGACGCCCTCTGTGGGTGGAGCTGCCCTCGGAGGGACGCCGGCTGGTA  
1321 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1380  
A G D A P S V G G S C R S E G T P A W Y -  
  
CATGCACGGCGAACGGGCCACTACTCCCACACCATGCCATGCGCAGCCCCAGCAAGCT  
1381 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1440  
M H G E R A H Y S H T M P M R S P S K L -  
  
CAGCCATATCTCCCGCTGGAGCTGGTGAATCCCTGGACTCGGATGAGGTGGACCTCAA  
1441 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1500  
S H I S R L E V E S L D S D E V D L K -  
  
GTCCGGCTACATGAGCGACAGTGACCTCATGGCAAGACCATGACGGAGGATGACAT  
1501 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1560  
S G Y M S D S D L M G K T M T E D D D I -  
  
.CACTACCGGCTGGGATGAAAGCAGCTCCATCAGTAGTGGACTCAGCGATGCCTCAGACAA  
1561 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1620  
T T G W D E S S S I S S G L S D A S D N -  
  
TCTCAGTTAGAAGAATTCAATGCCAGCTCCTCACTCAACTCCCTCCCAAGTACTCCCAC  
1621 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1680  
L S S E E F N A S S S L N S L P S T P T -  
  
TGCTTCTCGCAGGAACACTAACAAATAGTGCACAGACTCAGAGAAGCGCTCACTGGC  
1681 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1740  
A S R R N S T I V L R T D S E K R S L A -

Fig. 16 (cont'd 1)

09/914549

80/124

AGAAAGTGGCTGAGCTGGTTAGTGAATCAGAGGAGAAAGCCCCCTAAAAAACTGGAGTA  
 1741 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1800  
 E S G L S W F S E S E E K A P K K L E Y -  
  
 CGACAGTGGTAGCCTGAAGATGGAACCTGGACTCTAAGTGGCGGAGGGAGCGGCCTGA  
 1801 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1860  
 D S G S L K M E P G T S K W R R E R P E -  
  
 GAGCTGTGATTCAAGGGTGAGAACTGAAAAAGCCCAGCCTGGCCACCC  
 1861 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1920  
 S C D D S S K G G E L K K P I S L G H P -  
  
 TGGTCCCTGAAGAAGGGCAAGACCCCACCTGTGGCTGTAACCTCCCCATCACTCACAC  
 1921 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1980  
 G S L K K G K T P P V A V T S P I T H T -  
  
 AGCCCAGAGTGCCCTCAAAGTCGCAGGCAAACCTGAGGGCAAAGCTACAGACAAGGGTAA  
 1981 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2040  
 A Q S A L K V A G K P E G K A T D K G K -  
  
 GCTTGCAGTGAAGAATACTGGCTCCAACGCTCCTCTGATGCTGGCGGACCGCCT  
 2041 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2100  
 L A V K N T G L Q R S S S D A G R D R L -  
  
 GAGTGATGCTAAGAAGCCCCCTCGGGCATTGCTCGCCCTCCACTTCGGGATCCTTGG  
 2101 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2160  
 S D A K K P P S G I A R P S T S G S F G -  
  
 CTACAAGAAGCCTCCTGCCACAGGCACAGCCACTGTCATGCAAACCTGGTGGTCAGC  
 2161 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2220  
 Y K K P P P A T G T A T V M Q T G G S A -  
  
 CACTCTCAGCAAGATCCAGAAGTCCTCAGGCATCCCTGTCAAGCCAGTAAATGGCGCAA  
 2221 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2280  
 T L S K I Q K S S G I P V K P V N G R K -  
  
 GACTAGCTTAGATGTTCCAACAGTGCAGAGCCAGGATTCTGGCTCCTGGAGCCGTT  
 2281 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2340  
 T S L D V S N S A E P G F L A P G A R S -  
  
 TAACATCCAGTACCGCAGCCTGCCCGGCCAGCCAAGTCAAGTTCTATGAGCGTGACCGG  
 2341 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2400  
 N I Q Y R S L P R P A K S S S M S V T G -  
  
 CGGGCGGGGTGGACCTCGCCCTGTGAGCAGCAGCATTGACCCAGTCTCCTCAGCACCAA  
 2401 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2460  
 G R G G P R P V S S S I D P S L L S T K -  
  
 GCAGGGAGGCCTTACGCCTTCCAGACTGAAGGAGCCTACCAAGGTAGCCAGTGGCGGAC  
 2461 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2520  
 Q G G L T P S R L K E P T K V A S G R T -  
  
 CACTCCAGCCCCGTCAATCAGACAGATCGGGAAAAGGAGAAGGCCAAGCCAAGGAGT  
 2521 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2580  
 T P A P V N Q T D R E K E K A K A K A V -  
  
 GGCCTGGACTCAGACAACATCTCCTGAAGAGTATTGGCTCCAGAAAGTACTCCCAA  
 2581 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2640  
 A L D S D N I S L K S I G S P E S T P K -

Fig. 16 (cont'd 2)

GAACCAAGCAAGCCACCCCACAGCCACCAAGCTGGCAGAGCTGCCACCAACCCCTCTCAG  
 2641 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2700  
 N Q A S H P T A T K L A E L P P T P L R -  
  
 GGCCACAGCGAAGAGCTTGTCAAACCACCCCTCACTAGCCAATCTTGACAAGGTCAACTC  
 2701 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2760  
 A T A K S F V K P P S L A N L D K V N S -  
  
 CAACAGTCTGGATCTACCATCATCCAGTGATACCACCCATGCTCAAAGGTCCCAGATCT  
 2761 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2820  
 N S L D L P S S S D T T H A S K V P D L -  
  
 GCATGCTACAAGCTCAGCATCTGGGGCCCTCTCCCTCCTGCTCACCCCCAGTCCGGC  
 2821 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2880  
 H A T S S A S G G P L P S C F T P S P A -  
  
 ACCCATCCTCAATATTAACCTAGCCAGCTCTCCAGGGCCTGGAGCTAATGAGTGGTTT  
 2881 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2940  
 P I L N I N S A S F S Q G L E L M S G F -  
  
 CAGTGTGCCAAAAGAGACCCGCATGTACCCCAAACCTCTCAGGCCTGCACAGGAGCATGGA  
 2941 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3000  
 S V P K E T R M Y P K L S G L H R S M E -  
  
 GTCCCTCCAGATGCCAATGAGCCTCCCCAGTGCCTTCCCCAGCAGTACTCCGTCCCCAC  
 3001 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3060  
 S L Q M P M S L P S A F P S S T P V P T -  
  
 CCCACCTGCTCCCCCTGCTGCTCCCACAGAAGAAGAGACGGAAGAGCTGACTTGGAGTGG  
 3061 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3120  
 P P A P P A A P T E E E T E E L T W S G -  
  
 AAGCCCCAGAGCTGGCAACTGGACAGTAATCAGCGGGATCGGAACACTCTTCCCAAGAA  
 3121 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3180  
 S P R A G Q L D S N Q R D R N T L P K K -  
  
 AGGGCTCAGGTACCAGCTTCAGTCCCAGGAGGAGACCAAGGAGAGGCGACATTCCCATAAC  
 3181 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3240  
 G L R Y Q L Q S Q E E T K E R R H S H T -  
  
 CATTGGTGGCTGCCGAATCCGATGACCAGTCAGAGCTGCCTTCTCCCCCTGCACTTCC  
 3241 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3300  
 I G G L P E S D D Q S E L P S P P A L P -  
  
 CATGTCTCTGAGTGCCTAAAGGGCAACTTACCAACATAGTGAGTCCCCTGCCTGCCACCAC  
 3301 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3360  
 M S L S A K G Q L T N I V S P T A A T T -  
  
 \*GCCAAGAACATCCCCCTCCAAACAGCATTCCCCACCCACGAGGCGGGCTTCGAGCTGTACAG  
 3361 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3420  
 P R I T R S N S I P T H E A A F E L Y S -  
  
 CGGCTCCCAAATGGGGAGCACCCCTGTCCTGGCCGAGAGACCCAAGGGAATGATTGGTC  
 3421 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3480  
 G S Q M G S T L S L A E R P K G M I R S -  
  
 AGGATCCTTCCGAGACCCACGGACGATGTTACGGCTCAGTGCTGCTCCCTGGCCTCCAG  
 3481 - +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3540  
 G S F R D P T D D V H G S V L S L A S S -

Fig. 16 (cont'd 3)

82/124

TGCCTCCTCCACCTACTCCTCAGCTGAGGAGAGGATGCAATCTGAGCAAATCCGGAAGCT  
 3541 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3600  
 A S S T Y S S A E E R M Q S E Q I R K L -

TCGTAGGGAACTGGAATCATCCCAGGAAAAAGTGGCCACCTTGACGTCTCAGCTTCTGC  
 3601 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3660  
 R R E L E S S Q E K V A T L T S Q L S A -

CAATGCTAATCTGGTGGCTGTTTGAGCAGAGCCTGGTGAATATGACATCCGCCCTGCG  
 3661 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3720  
 N A N L V A A F E Q S L V N M T S R L R -

ACACCTGGCAGAGACGGCCGAGGAGAAGGACACTGAGCTGCTGGATTGCGAGAAACCAT  
 3721 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3780  
 H L A E T A E E K D T E L L D L R E T I -

AGACTTTCTGAAGAAAAAGAACCTTGAGGCCAGGCAGTCATTAGGGAGCCCTTAATGC  
 3781 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3840  
 D F L K K K N S E A Q A V I Q G A L N A -

CTCAGAAACCACACCCAAAGAACCTCGGATCAAGAGACAAACTCCTCAGATAGCATCTC  
 3841 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3900  
 S E T T P K E L R I K R Q N S S D S I S -

AAGCCTAACAGCATCACTAGCCATTCCAGCATCGGAGCAGCAAGGATGCTGATGCGAA  
 3901 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3960  
 S L N S I T S H S S I G S S K D A D A K -

AAAGAAGAAAAAAAAGAGTTGGCTTCGAAGTTCTCAACAAAGCGTTCAAGTATAAAAAAA  
 3961 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4020  
 K K K K K S W L R S S F N K A F S I K K -

GGGGCCAAGTCAGCTTCTCATACTCGGATATAGAGGAGATTGCTACACCCGACTCTC  
 4021 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4080  
 G P K S A S S Y S D I E E I A T P D S S -

AGCCCCCTCATCCCCAAACTACAGCATGGTCTACAGAGACTGCTCACCCCTCATCAA  
 4081 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4140  
 A P S S P K L Q H G S T E T A S P S I K -

GTCCTCCACCTCGTCCCTGGCACTGATGTACCGAGGGCCCTGCTCACCCAGCCCC  
 4141 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4200  
 S S T S S S V G T D V T E G P A H P A P -

CCACACTAGGCTGTTCCATGCAAATGAGGAGGAGGCCAGAGAAGAAGGAGGTATCGGA  
 4201 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4260  
 H T R L F H A N E E E E P E K K E V S E -

GCTCGCTCTGAGCTATGGGAGAAGGAAATGAAGCTTACAGACATCCGTTGGAGGCCCT  
 4261 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4320  
 L R S E L W E K E M K L T D I R L E A L -

CAAECTGCCCACCAACTGGATCAGCTGGAGACCATGCACAACATGCAGTTGGAGGT  
 4321 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4380  
 N S A H Q L D Q L R E T M H N M Q L E V -

GGACCTGCTGGAAGCAGAGAATGACCGACTGAAGGTAGCCCCAGGCCCTCATCAGGCTC  
 4381 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4440  
 D L L E A E N D R L K V A P G P S S G S -

Fig. 16 (cont'd 4)

83/124

CACTCCAGGGCAGGTCCCTGGATCATCTGCATTATCTTCCCCACGCCGCTCCCTAGGCCT 4500  
 4441 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 T P G Q V P G S S A L S S P R R R S L G L -  
  
 GGCACTCACCCATTCTCGGCCCCAGTCTTGAGACACAGACCTGTCACCCATGGATGG 4560  
 4501 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 A L T H S F G P S L A D T D L S P M D G -  
  
 CATCAGTACTTGTGGTCCAAAGGGAGGAAGTGACCCCTCCGGTGGTGGTGGAGGATGCC 4620  
 4561 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 I S T C G P K E E V T L R V V V R M P P -  
  
 GCAGCACATCATCAAAGGGACTTGAAGCAGCAGGAATTCTTCCTGGGCTGTAGCAAGGT 4680  
 4621 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 Q H I I K G D L K Q Q E F F L G C S K V -  
  
 CAGTGGAAAAGTTGACTGGAAGATGCTGGATGAAGCTGTTTCCAAGTGTCAAGGACTA 4740  
 4681 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 S G K V D W K M L D E A V F Q V F K D Y -  
  
 TATTTCTAAAATGGACCCAGCCTCTACCCCTGGACTAACGACTGAGTCATCCATGGCTA 4800  
 4741 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 I S K M D P A S T L G L S T E S I H G Y -  
  
 CAGCATCAGCCACGTGAAACGAGTGGATGCAGAGCCCCCGAGATGCCTCCTGCCG 4860  
 4801 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 S I S H V K R V L D A E P P E M P P C R -  
  
 TCGAGGTGTCAATAACATATCAGTCTCCCTCAAAGGTCTGAAGGAGAAATGCGTCGACAG 4920  
 4861 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 R G V N N I S V S L K G L K E K C V D S -  
  
 CCTGGTGGCTGAGACGCTGATCCCCAAGCCGATGATGCAGCACTACATAAGCCTCCTGCT 4980  
 4921 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 L V F E T L I P K P M M Q H Y I S L L L -  
  
 GAAGCACCGGGCGCCTCGTCCCTCGGGCCCCAGCGGCACGGCAAGACCTACCTGACCAA 5040  
 4981 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 K H R R L V L S G P S G T G K T Y L T N -  
  
 TCGCTTGGCCGAGTACCTGGGGAGCGCTCTGGCGTGAGGTACAGAGGGCATCGTCAG 5100  
 5041 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 R L A E Y L V E R S G R E V T E G I V S -  
  
 CACCTTCAACATGCACCAGCAGTCTTCAAGGATCTGCAACTGTATCTTCAACCTAGC 5160  
 5101 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 T F N M H Q Q S C K D L Q L Y L S N L A -  
  
 CAACCAAGATAGACCGGGAAACAGGAATTGGGGATGTGCCCTGGTATTCTATTGGATGA 5220  
 5161 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 N Q I D R E T G I G D V P L V I L L D D -  
  
 CCTGAGTGAAGCAGGCTCCATCAGTGAGTTGGTCAATGGGGCCCTCACCTGCAAGTATCA 5280  
 5221 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 L S E A G S I S E L V N G A L T C K Y H -  
  
 TAAATGTCCCTATATTATAGGTACCACCAATCAGCCTGTAAAAATGACACCCAACCATGG 5340  
 5281 - +-----+-----+-----+-----+-----+-----+-----+-----+  
 K C P Y I I G T T N Q P V K M T P N H G -

Fig. 16 (cont'd 5)

84/124

CTTCACTTGAGCTTCAGGATGTTGACCTCTCCAACAACAGTGGAGCCAGCCAATGGCTT  
5341 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5400  
F H L S F R M L T F S N N V E P A N G F -

CCTGGTTCGTTACCTGAGGAGGAAGCTGGTAGAGTCAGACAGCGACATCAATGCCAACAA  
5401 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5460  
L V R Y L R R K L V E S D S D I N A N K -

GGAAGAGCTGCTTCGGGTGCTCGACTGGGTACCCAAGCTGTGGTATCATCTCCACACCTT  
5461 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5520  
E E L L R V L D W V P K L W Y H L H T F -

CCTTGAGAAGCACAGCACCTCAGACTTCCTCATCGGCCCTGCTCTTCTGCGTGTCC  
5521 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5580  
L E K H S T S D F L I G P C F F L S C P -

CATTGGCATTGAGGACTTCGGACCTGGTTCATTGACCTGTGGAACAACCTATCATTC  
5581 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5640  
I G I E D F R T W F I D L W N N S I I P -

CTATCTACAGGAAGGAGCCAAGGATGGATAAAGGTCCATGGACAGAAAGCTGCTGGGA  
5641 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5700  
Y L Q E G A K D G I K V H G Q K A A W E -

GGACCCAGTGGATGGTCCGGACACACTCCCTGGCATCAGCCAACAGACCAATC  
5701 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5760  
D P V E W V R D T L P W P S A Q Q D Q S -

AAAGCTGTACCACCTGCCCAACCCACCGTGGCCCTCACAGCATTGCCACCTCCGA  
5761 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5820  
K L Y H L P P P T V G P H S I A S P P E -

GGATAGGACAGTCAAAGACAGCACCCCAAGTTCTCTGGACTCAGATCCTCTGATGGCCAT  
5821 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5880  
D R T V K D S T P S S L D S D P L M A M -

GCTGCTGAAACTTCAAGAAGCTGCCACTACATTGAGTCTCCAGATCGAGAACCATCCT  
5881 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5940  
L L K L Q E A A N Y I E S P D R E T I L -

GGACCCCAACCTTCAGGCAAACACTTAAGGGTTCGGCAATCACTGTCACCCCGGACAGC  
5941 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6000  
D P N L Q A T L \*

AGAACGCTGGCATCAGCTATCTAGCTCCTCTCCCTCTCAGAGCACTGG  
6001 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6060

CTCTCCAGCCCCAGGAGGAGAACAGGAGGGAGGAGATGAAAGAGGAGGGACAGGTT  
6061 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6120

TTGGTGCTGTACCTTGAGAACCTCCTAGGAAGGAATGGTGGGTGGCGTTGGAACTT  
6121 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6180

GTCCCCCTAAACACATTACTGGCCTCCTAATGACTTGGGGAAAAGATGATTCTGG  
6181 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6240

GTCTTCCCTTGACTTCTGTTCAATTACAAACTCCTGGCTTCTGGGGAGGGGTTCA  
6241 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6300

Fig. 16 (cont'd 6)

09/914549

85/124

GAAAACATCAAAACACTGCAGCAGTCTCAAATGATTCTCACAGCAACCCTGAGAGAGA  
6301 -----+-----+-----+-----+-----+-----+ 6360

CAGTCTTGTGAGGGAGATCTGGGGAGGCAGGAAGCTCCTCAGATTTCTCACAGACCCT  
6361 -----+-----+-----+-----+-----+-----+ 6420

TCCAATTCCATCACCACTGCCAACAACTCCTCCCCAGAGATCTGGCTGGAGCCCAGAA  
6421 -----+-----+-----+-----+-----+-----+ 6480

AAAGAAGCATGTGGTTAAAAAATGTTAAATCAATCTGTAAAAGGTAAAATGAAAAAC  
6481 -----+-----+-----+-----+-----+-----+ 6540

AAAAACAAGCAAACAAACAAAAACAATGGAAAAGATGAAGCTGGAGAGAGAGGAACCAG  
6541 -----+-----+-----+-----+-----+-----+ 6600

TTGCCAAGGTAGAGAGCTGCCGCTCCTGCCCTCTGGATGACATAGGGACATCAACAAG  
6601 -----+-----+-----+-----+-----+-----+ 6660

ACGGCTGCCAACCTGAGAAGTCACCAACCACAAAAATAACCTTACAGCCTTCAGGGAAA  
6661 -----+-----+-----+-----+-----+-----+ 6720

GACTACCAGCTCTGTCTTCTACCCCTTAATTAAACAATGCATAAGAGTCAATAACCCT  
6721 -----+-----+-----+-----+-----+-----+ 6780

ACTTTTTAAAAAAAAAAAAAG  
6781 -----+-----+----- 6805

Fig. 16 (cont'd 7)

T3-cDNA sequence and T3 protein  
encoded therein (protein isoform 1)

```

CAACCAGCCAGAACGCCCTGAACTCGCAGGTGCTGCAGGGCTGCAGGAGCCAGCGGGGAA
1   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 60
N Q P E R L N S Q V L Q G L Q E P A G E -
GGGGCTCCCGCTGCCGAAGAGCGGCTCGTGGAAAACGGGTCGATACCCAGATCTACAC
61  -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 120
G L P L R K S G S V E N G F D T Q I Y T -
AGACTGGCCAATCATTACCTAGCAAATCCGCCACAAGCGTCTCATCAGGGATCTCCA
121 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 180
D W A N H Y L A K S G H K R L I R D L Q -
GCAAGATGTGACAGATGGCTGCCCTGGCCCAGATTATCCAGGTTGTGGCAAATGAAAAA
181 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 240
Q D V T D G V L L A Q I I Q V V A N E K -
GATTGAAGACATCAATGGCTGCCGAAGAACAGATCCCAAATGATTGAAAACATAGATGC
241 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 300
I E D I N G C P K N R S Q M I E N I D A -
CTGCTTGAATTCTGGCAGCTAACGGAAATAAACATCCAGGGCTGTCTGCAGAACAGAT
301 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 360
C L N F L A A K G I N I Q G L S A E E I -
CAGGAATGGAAACCTCAAGGCCATTCTAGGCCTCTCAGCCTCTCCGATACAAGCA
361 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 420
R N G N L K A I L G L F F S L S R Y K Q -
GCAGCAGCAGCAGCCCCAGAACAGCACCTCTCCTCACCTCTGCCGCCGTATCCCA
421 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 480
Q Q Q Q P Q K Q H L S S P L P P A V S Q -
GGTGGCCGGGGCCCCCTCCAGTGCCAGGCTGGCACCCCTCAGCAGCAGGTGCCAGTCAC
481 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 540
V A G A P S Q C Q A G T P Q Q Q V P V T -
TCCCCAAGCCCCGTGCCAGCCTCACCAAGCCAGCAGGCCACATCAGCAGTCAGCACAAGC
541 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 600
P Q A P C Q P H Q P A P H Q Q S K A Q A -
TGAAATGCAGTCCAGACTTCCAGGTCTACCGCGAGGGTATCCGCTGCAGGCAGCGAGGC
601 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 660
E M Q S R L P G P T A R V S A A G S E A -
CAAAACACCGCGGAGGGTCAACTACTGCTAACAAACCGACGCAGCCAGAGCTTTAACAACTA
661 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 720
K T R G G S T T A N N R R S Q S F N N Y -
TGATAAAATCCAAACCAGTCACCTCCCCACCCCAACGCCAACGCAGCCACGAGAAAGAGCC
721 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 780
D K S K P V T S P P P P S S H E K E P -
TTTGGCAAGTTCAGCCTCTCCACCCCGAATGAGTGACAATGCACCTGCTTCCTTGGAA
781 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 840
L A S S A S S H P G M S D N A P A S L E -

```

Fig. 17

09/914549

87/124

GAGCGGCAGCAGCTCCACCCCTACTAATTGAGTACCTCCTGGCCATCCCGCAGCCCGG  
 841 - +-----+-----+-----+-----+-----+-----+-----+-----+ 900  
 S G S S S T P T N C S T S S A I P Q P G -  
  
 TGCAGCCACCAAGCCTTGGCGCAGCAAATCCCTCAGCGTGAAGCACAGTGCCACGGTATC  
 901 - +-----+-----+-----+-----+-----+-----+-----+-----+ 960  
 A A T K P W R S K S L S V K H S A T V S -  
  
 CATGCTCTCGGTCAAGCCTCTGGCCTGAGGCCACACCTGAAGCCATGAA  
 961 - +-----+-----+-----+-----+-----+-----+-----+-----+ 1020  
 M L S V K P P G P E A P R P T P E A M K -  
  
 GCCGGCCCCAACAAATCAGAAGTCATGCTGGAAAAGCTGAAACTTTAACAGTAAAGG  
 1021 - +-----+-----+-----+-----+-----+-----+-----+-----+ 1080  
 P A P N N Q K S M L E K L K L F N S K G -  
  
 GGGCTCAAAGGCAGGTGAGGGGCCGGGTCCGGACACAAGCTGTGAGCGGCTGGAGAC  
 1081 - +-----+-----+-----+-----+-----+-----+-----+-----+ 1140  
 G S K A G E G P G S R D T S C E R L E T -  
  
 TCTGCCAGCTTCGAAGAGAGCGAGGAGCTGGAGGCCAGTCATGCTCACCAACCGT  
 1141 - +-----+-----+-----+-----+-----+-----+-----+-----+ 1200  
 L P S F E E S E E L E A A S R M L T T V -  
  
 GGGCCCTGCTTCCAGCAGCCCCAAGATTGCACTCAAGGGATTGCCAGAGGACTTTAG  
 1201 - +-----+-----+-----+-----+-----+-----+-----+-----+ 1260  
 G P A S S S P K I A L K G I A Q R T F S -  
  
 CCGGGCACTGACCAACAAGAAGAGTTCTCTGAAAGGCAATGAGAAAGAGAAGGAGAAACA  
 1261 - +-----+-----+-----+-----+-----+-----+-----+-----+ 1320  
 R A L T N K K S S L K G N E K E K E K Q -  
  
 ACAGCGGGAGAAGGATAAGGAGAAAAGCAAGGACCTTGCCAAGAGAGCCTCTGTGACGGA  
 1321 - +-----+-----+-----+-----+-----+-----+-----+-----+ 1380  
 Q R E K D K E K S K D L A K R A S V T E -  
  
 GAGGCTGGACCTCAAGGAGGCCAAAAGAAGACCCAGTGGAGCAGCTGTGCCAGAT  
 1381 - +-----+-----+-----+-----+-----+-----+-----+-----+ 1440  
 R L D L K E E P K E D P S G A A V P E M -  
  
 GCCAAAAAAAGTCCTCAAGATTGCCAGCTTCATCCCCAAGGGGGAGCTAACAGTGC  
 1441 - +-----+-----+-----+-----+-----+-----+-----+-----+ 1500  
 P K K S S K I A S F I P K G G K L N S A -  
  
 CAAGAAGGAGCCCAGGCCCTTCCCACAGTGGAAATACCAAAACCAGGAATGAAGAGCAT  
 1501 - +-----+-----+-----+-----+-----+-----+-----+-----+ 1560  
 K K E P M A P S H S G I P K P G M K S M -  
  
 GCCCGGGAAATCCCCAAGTGGCCAGCGCCTTCCAAGGAAGGGAGCGGAGCGGAGTGG  
 1561 - +-----+-----+-----+-----+-----+-----+-----+-----+ 1620  
 P G K S P S A P A P S K E G E R S R S G -  
  
 GAAGCTGAGCTCAGGACTCCCCCAGCAGAAGCCCCAGTGGACGGCAGACACTCCAGTTC  
 1621 - +-----+-----+-----+-----+-----+-----+-----+-----+ 1680  
 K L S S G L P Q Q K P Q L D G R H S S S -  
  
 CTCTTCCAGCCTGGCGTCCCTCAGAAGGAAAAGGCCAGGAGGGACCACCTGAACCACAG  
 1681 - +-----+-----+-----+-----+-----+-----+-----+-----+ 1740  
 S S S L A S S E G K G P G G T T L N H S -

Fig. 17 (cont'd 1)

88/124

CATCAGCAGCCAGACTGTCAGTGGCTGTCGGGACCACCCAGACCACAGGAAGCAATAC  
 1741 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1800  
 I S S Q T V S G S V G T T Q T T G S N T -  
  
 CGTCAGTGTTCAGCTACCTCAGCCCCAGCAGCAATACAACCATCCAAACACTGCCACGGT  
 1801 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1860  
 V S V Q L P Q P Q Q Q Y N H P N T A T V -  
  
 TGCACCTTCCTGTACAGGTCTCAGACGGACACTGAAGGGAATGTTACTGCCGAGTCAG  
 1861 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1920  
 A P F L Y R S Q T D T E G N V T A E S S -  
  
 CTCAACAGGTGTGAGCGTGGAGGCCAGCCACTCACCAAGACTGGACAGCCTGCTCTGGA  
 1921 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1980  
 S T G V S V E P S H F T K T G Q P A L E -  
  
 AGAACTCACTGGGAAGATCCTGAGGCTCGGCGGCTGGGACAGTGAAAGAACATCGCTGA  
 1981 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2040  
 E L T G E D P E A R R L R T V K N I A D -  
  
 TCTGCGGCAGAATTGGAGGAAACCATGTCCAGTTAAGGGAACTCAGGTTACACACAG  
 2041 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2100  
 L R Q N L E E T M S S L R G T Q V T H S -  
  
 CACATTGAAACCACGTTGACACCAATGTCACCAACGGAGATGAGTGGCCGTAGCATACT  
 2101 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2160  
 T L E T T F D T N V T T E M S G R S I L -  
  
 CAGCTTGACAGGGAGGCCACACCTCTGTCCCTGGAGACTGGGCCAGTCCAGCCCTGGCT  
 2161 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2220  
 S L T G R P T P L S W R L G Q S S P R L -  
  
 CCAAGCAGGAGACGCCCTCAATGGCAATGGTATCCCCCTCGAGCCAACGCCAGCAG  
 2221 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2280  
 Q A G D A P S M G N G Y P P R A N A S R -  
  
 GTTCATCAACACTGAGTCAGTCGCTATGTGTACTCCGCCCTCTGAGAAGGCAGCTGGC  
 2281 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2340  
 F I N T E S G R Y V Y S A P L R R Q L A -  
  
 CTCCCGGGCAGTAGTGTCTGCCACGTGGACGTCTCAGACAAGGCAGGAGATGAGATGGA  
 2341 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2400  
 S R G S S V C H V D V S D K A G D E M D -  
  
 CCTGGAAGGCATCAGCATGGACGCCCGGCTACATGAGCGATGGGATGTTCTGAGCAA  
 2401 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2460  
 L E G I S M D A P G Y M S D G D V L S K -  
  
 GAACATCCGGACCGATGACATTACAAGCGGATACATGACTGATGGTGGACTTGGCTCTA  
 2461 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2520  
 N I R T D D I T S G Y M T D G G L G L Y -  
  
 TACCCGTCGCCTGAACCGGCTCCCTGATGGATGGCTGTGGTACGGGAGACCCTGCAACG  
 2521 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2580  
 T R R L N R L P D G M A V V R E T L Q R -  
  
 AAATACCTCCCTGGCCTCGGAGACGCTGACAGCTGGACACAGCAGCTCCGTACAGCAG  
 2581 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2640  
 N T S L G L G D A D S W D D S S S S V S S -

Fig. 17 : (cont'd 2)

CGGCATCAGCGACACCATA GACAACCTCAGCACTGATGACATCAACACCAGCTCCCAT  
 2641 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2700  
 G I S D T I D N L S T D D I N T S S S S I -  
  
 CAGCTCTTATGCCAACACACCTGCCTCTCGAAAAACCTGGATGTGCAGACTGATGC  
 2701 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2760  
 S S Y A N T P A S S R K N L D V Q T D A -  
  
 TGAGAACGCACTCACAGGTGGAGAGGAATTCCCTGTGGTCTGGTATGATGTCAAGAAC  
 2761 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2820  
 E K H S Q V E R N S L W S G D D D V K K S -  
  
 AGACGGAGGCTCAGACAGCGGCATAAAATGGAGCCAGGTTCCAAGTGAGGCGGAATCC  
 2821 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2880  
 D G G S D S G I K M E P G S K W R R N P -  
  
 TTCTGATGTGCTGACGAGTCCGACAAAAGCACGTCGGCAAGAAGAATCCTGTCATCTC  
 2881 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2940  
 S D V S D E S D K S T S G K K N P V I S -  
  
 CCAGACAGGCTCATGGCGCGAGGCATGACAGCTCAGGTGGCATCACCAGGCCAAGGAC  
 2941 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3000  
 Q T G S W R R G M T A Q V G I T M P R T -  
  
 GAAGGCTTCAGCCCCGGCAGGCGACTGAAGACCCCAGGAACCTGGAAAAACAGACGACGC  
 3001 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3060  
 K A S A P A G A L K T P G T G K T D D A -  
  
 AAAGGTGTCTGAGAAAGGAAGGCTTCTCTAAAGCCTCCCAGGTGAAGCGCTCCCCATC  
 3061 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3120  
 K V S E K G R L S P K A S Q V K R S P S -  
  
 AGATGCAGGCCGGAGCAGTGGTGACGAATCCAAAAAGCCCCCTCCCCAGCAGCTCTAGGAC  
 3121 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3180  
 D A G R S S G D E S K K P L P S S S R T -  
  
 ACCTACTGCCAATGCCAACAGCTTGGTTCAAGAACAGCAGAGTGGTCCGCCACCGGCCT  
 3181 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3240  
 P T A N A N S F G F K K Q S G S A T G L -  
  
 GGCCATGATCACAGCCAGCGGGTGACTGTCACCAGCAGGTCAAGCCACACTGGCAAAAT  
 3241 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3300  
 A M I T A S G V T V T S R S A T L G K I -  
  
 CCCAAAGTCATCTGCACTCGTCAGTCGGTCTGCTGGTCGGAAGTCAGTATGGATGGGC  
 3301 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3360  
 P K S S A L V S R S A G R K S S M D G A -  
  
 TCAGAACAGGATGACGGGTATCTAGCCCTAACAGCTCCGGACAAACCTTCAGTACCGGAG  
 3361 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3420  
 Q N Q D D G Y L A L S S R T N L Q Y R S -  
  
 TTTGCCGAGGCCAGTAAGTCCAACAGCCGGAACGGGCTGGAACAGGTCTAGCACCAG  
 3421 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3480  
 L P R P S K S N S R N G A G N R S S T S -  
  
 CAGCATAGATTCCAACATTAGCAGCAAGTCCGCAGGCCTGCCAGTGCCAAACTGAGGGA  
 3481 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3540  
 S I D S N I S S K S A G L P V P K L R E -

Fig. 17 (cont'd 3)

09/914549

90/124

GCCTTCCAAAACAGCCCTAGGCAGCTCTCTACCAGGTCTGGTCAACCAAACAGACAAGGA  
 3541 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3600  
 P S K T A L G S S L P G L V N Q T D K E -

GAAAGGCATCTCATCAGACAACGAGAGTGTGGCTCCTGTAACTCGGTGAAAGTGAATCC  
 3601 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3660  
 K G I S S D N E S V A S C N S V K V N P -

GGCAGCCCAGCCTGTGTCCAGTCCGGCTCAGACCAAGTCTCCAGCCTGGAGCCAAGTACCC  
 3661 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3720  
 A A Q P V S S P A Q T S L Q P G A K Y P -

AGATGTGGCCTCTCCCACACTCCGCAGACTCTTGGTGGGAAGCCTACCAAGCAAGTGCC  
 3721 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3780  
 D V A S P T L R R L F G G K P T K Q V P -

CATGCCACAGCTGAAAACATGAAAATTGGTGGTCATCTCAATCCTCATGCCACCAT  
 3781 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3840  
 I A T A E N M K N S V V I S N P H A T M -

GACTCAGCAAGGTAACCTAGACTCCCCGTAGGCAGTGGCGCTCTGAGCAGTGGGAGCAG  
 3841 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3900  
 T Q Q G N L D S P S G S G V L S S G S S -

CAGTCCTCTCTACAGCAAGAATGTGGACCTCAACCAGTCTCCGCTAGCCTCCAGCCCCAG  
 3901 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3960  
 S P L Y S K N V D L N Q S P L A S S P S -

CTCAGCCCACCGGCCCTTCAACAGCCTCACCTGGGGCACCAACGCCAGCAGCTCCTC  
 3961 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4020  
 S A H S A P S N S L T W G T N A S S S S -

CGCAGTTAGCAAGGATGGCTGGCTTCACTGTCAGCAGCCTCCACACCAGCTGTGA  
 4021 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4080  
 A V S K D G L G F Q S V S S L H T S C E -

GTCCATCGACATCTCCCTCAGCAGTGGAGGGTCCCCAGCCACAATTCTTCACTGGCCT  
 4081 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4140  
 S I D I S L S S G G V P S H N S S T G L -

CATCGCCTCTCCAAGGACGACTCCTTGACTCCCTTGTCAAGAACTAACAGTGTGAAGAC  
 4141 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4200  
 I A S S K D D S L T P F V R T N S V K T -

CACACTGTCAGAAAGCCCTCTCTCTTCCCTGCTGCTAGCCCTAAGTCTGCAGAAGTAC  
 4201 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4260  
 T L S E S P L S S P A A S P K F C R S T -

TCTGCCAGGAAACAGGACAGTGACCCGCACCTTGATAGGAACACTTGCCTAAGAAAG  
 4261 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4320  
 L P R K Q D S D P H L D R N T L P K K G -

ACTCAGGTATACTCCCACCTCCAGCTCGCACGCAAGAAGATGCAAAAGAATGTTACG  
 4321 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4380  
 L R Y T P T S Q L R T Q E D A K E W L R -

GTCCCATTCTGCAGGAGGCCTTCAGGACACCGCTGCCAATTCCCCCTTTCTGGCTC  
 4381 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4440  
 S H S A G G L Q D T A A N S P F S S G S -

Fig. 17 (cont'd 4)

CAGCGTACTTCTCCCTCCGGAACAGATTCAACCTTTCCCAGCTTGCGAGTCCCACAC  
 4441 -----+-----+-----+-----+-----+-----+-----+-----+ 4500  
 S V T S P S G T R F N F S Q L A S P T T -  
  
 TGTCACCCAGATGAGCTTGTCCAACCCGACCATGCTGAGGACTCACAGCCTCTCCAATGC  
 4501 -----+-----+-----+-----+-----+-----+-----+-----+ 4560  
 V T Q M S L S N P T M L R T H S L S N A -  
  
 TGATGGGCAGTATGATCCATACTGACAGCCGCTTCCGGAAATAGCTCCATGTCCCTGGA  
 4561 -----+-----+-----+-----+-----+-----+-----+-----+ 4620  
 D G Q Y D P Y T D S R F R N S S M S L D -  
  
 TGAGAAGAGCAGAACCATGAGCCGTTCAGGCTCATTCCGGATGGGTTGAAGAAGTTCA  
 4621 -----+-----+-----+-----+-----+-----+-----+-----+ 4680  
 E K S R T M S R S G S F R D G F E E V H -  
  
 TGGATCCTCACTTCCTGGTTCCAGCACATCGTCAGTTATTCTACACCAAGAAGAAAA  
 4681 -----+-----+-----+-----+-----+-----+-----+-----+ 4740  
 G S S L S L V S S T S S V Y S T P E E K -  
  
 ATGCCAGTCAGAGATTGCAAGCTGCGCGGAACTGGATGCCTCCCAGGAGAAAGTTTC  
 4741 -----+-----+-----+-----+-----+-----+-----+-----+ 4800  
 C Q S E I R K L R R E L D A S Q E K V S -  
  
 AGCTTTGACCACCCAGCTGACAGCAAATGCTCACCTTGTGGCTGCCTTGAAACAGAGTCT  
 4801 -----+-----+-----+-----+-----+-----+-----+-----+ 4860  
 A L T T Q L T A N A H L V A A F E Q S L -  
  
 TGGTAACATGACAATCAGGCTCCAGAGTCTGACCATGACAGCTGAGCAGAAGGATTCA  
 4861 -----+-----+-----+-----+-----+-----+-----+-----+ 4920  
 G N M T I R L Q S L T M T A E Q K D S E -  
  
 ACTGAATGAGTTAAGAAAAACCATTGAGCTGCTAAAGAAACAGAACGAGCTGCCAGGC  
 4921 -----+-----+-----+-----+-----+-----+-----+-----+ 4980  
 L N E L R K T I E L L K K Q N A A A Q A -  
  
 TGCCATTAATGGAGTAATTAACACACCTGAGCTCAACTGCAAAGGAAACGGCACTGCCA  
 4981 -----+-----+-----+-----+-----+-----+-----+-----+ 5040  
 A I N G V I N T P E L N C K G N G T A Q -  
  
 GTCTGCAGACCTCCGCATCCGCAGGCAGCACTCCTCAGACAGCGTCTCCAGCATCACAG  
 5041 -----+-----+-----+-----+-----+-----+-----+-----+ 5100  
 S A D L R I R R Q H S S D S V S S I N S -  
  
 TGCCACCAGCCACTCCAGTGTGGGCAGCAACATAGAGAGTGAUTCAAAGAAGAAGAG  
 5101 -----+-----+-----+-----+-----+-----+-----+-----+ 5160  
 A T S H S S V G S N I E S D S K K K K R -  
  
 GAAGAACTGGGTCAATGAGTTACGCAGCTCCTCAAGCAAGCTTCTGGGAAGAAGAAGTC  
 5161 -----+-----+-----+-----+-----+-----+-----+-----+ 5220  
 K N W V N E L R S S F K Q A F G K K K S -  
  
 CCCAAAATCTGCGTCCTCTCATTAGATATTGAGGAGATGACGGATTCTTCTTGCCTTC  
 5221 -----+-----+-----+-----+-----+-----+-----+-----+ 5280  
 P K S A S S H S D I E E M T D S S L P S -  
  
 CTCACCAAAGTTACCAACACAATGGGTCCACAGGTTCCACCCACTGCTGAGGAATTCTCA  
 5281 -----+-----+-----+-----+-----+-----+-----+-----+ 5340  
 S P K L P H N G S T G S T P L L R N S H -

Fig. 17 (cont'd 5)

CTCCAACCTCTAAATTCAAGAATGCATGGGATAGTGAAGCTGAGACCGTCATGCAGCTCCG  
 5341 -----+-----+-----+-----+-----+-----+-----+-----+ 5400  
 S N S L I S E C M D S E A E T V M Q L R -  
  
 AAATGAGTTAACAGACAAGGAGATGAAGCTGACAGATATCCGCTTAGAAGCTCTCAGTTC  
 5401 -----+-----+-----+-----+-----+-----+-----+-----+ 5460  
 N E L R D K E M K L T D I R L E A L S S -  
  
 TGCCCACCAAGCTGGACCAGCTCCGGAGGCCATAACAGGATGCAGAGTGAAATAGAGAA  
 5461 -----+-----+-----+-----+-----+-----+-----+-----+ 5520  
 A H Q L D Q L R E A M N R M Q S E I E K -  
  
 GCTGAAAGCTGAGAACATGATCGGCTGAAGTCAGAGTCTCAAGGCAGTGGCTGCAGCCGGC  
 5521 -----+-----+-----+-----+-----+-----+-----+-----+ 5580  
 L K A E N D R L K S E S Q G S G C S R A -  
  
 TCCTTCCAAGTGTCCATCTGCCTCCCCGAGGCAGTCCATGGGCCTCTCCCAGCACAG  
 5581 -----+-----+-----+-----+-----+-----+-----+-----+ 5640  
 P S Q V S I S A S P R Q S M G L S Q H S -  
  
 CTTGAACCTCACTGAGTCACCGCCTGGACATGTTGCTGGATGACACTGGTGAATGCTC  
 5641 -----+-----+-----+-----+-----+-----+-----+-----+ 5700  
 L N L T E S T S L D M L L D D T G E C S -  
  
 GGCTCGGAAGGAAGGAGGCAGGCATGTTAACGATAGTTGTCAGCTTCAGGAGGAAATGAA  
 5701 -----+-----+-----+-----+-----+-----+-----+-----+ 5760  
 A R K E G G R H V K I V V S F Q E E M K -  
  
 GTGGAAGGAGGATTCCAGACCATCTCTTCTTATTGGCTGCATTGGAGTTAGTGGCAA  
 5761 -----+-----+-----+-----+-----+-----+-----+-----+ 5820  
 W K E D S R P H L F L I G C I G V S G K -  
  
 GACGAAGTGGATGTGCTCGATGGGTGGTAGACGGCTGTCAGGAAATACATCATTCA  
 5821 -----+-----+-----+-----+-----+-----+-----+-----+ 5880  
 T K W D V L D G V V R R L F K E Y I I H -  
  
 TGTCGACCCAGTGAGTCAGCTAGGGCTGAATTCAAGACAGCGTCTGGCTACAGCATTGG  
 5881 -----+-----+-----+-----+-----+-----+-----+-----+ 5940  
 V D P V S Q L G L N S D S V L G Y S I G -  
  
 AGAAAATCAAGCGCAGCAACACTTCCGAAACACCGGAGCTGCTTCCTGGCTATCTGGT  
 5941 -----+-----+-----+-----+-----+-----+-----+-----+ 6000  
 E I K R S N T S E T P E L L P C G Y L V -  
  
 TGGAGAGAACACGACCATCTCAGTGACTGTGAAAGGGCTCGCAGAAAACAGCCTGGACTC  
 6001 -----+-----+-----+-----+-----+-----+-----+-----+ 6060  
 G E N T T I S V T V K G L A E N S L D S -  
  
 ACTGGTGTGAGTCCTGATTCCCAAGCCCACCTGCAGCGCTACGTCTCCCTCCTGAT  
 6061 -----+-----+-----+-----+-----+-----+-----+-----+ 6120  
 L V F E S L I P K P I L Q R Y V S L L I -  
  
 AGAGCACCCTGGATCATTCTCTGGCCCCAGCGGACTGGAAAACCTACCTGGCAA  
 6121 -----+-----+-----+-----+-----+-----+-----+-----+ 6180  
 E H R R I I L S G P S G T G K T Y L A N -  
  
 CGGGCTGTCTGAGTATAGTGCCTCGAGAGGGACGGAGTTGACAGACGGGTTATCGC  
 6181 -----+-----+-----+-----+-----+-----+-----+-----+ 6240  
 R L S E Y I V L R E G R E L T D G V F A -

Fig. 17 (cont'd 6)

09/914549

93/124

CACCTTTAACGTGGACCATAAGTCCAGCAAGGAATTGCGCCAGTACCTGTCCAACCTTGC  
 6241 -----+-----+-----+-----+-----+-----+-----+ 6300  
 T F N V D H K S S K E L R Q Y L S N L A -

TGACCAGTGCAACAGT GAGAACAATGCTGTGGACATGCCCTCGTCATCATCCTGGACAA  
 6301 -----+-----+-----+-----+-----+-----+-----+ 6360  
 D Q C N S E N N A V D M P L V I I L D N -

CCTACACCACGTGAGCTCTGGCGAGATCTTCAATGGGCTGCTCAACTGCAAGTACCA  
 6361 -----+-----+-----+-----+-----+-----+-----+ 6420  
 GGATGTGGTGCACTCGAGAGACCCGCTCTAGAAGTTACCCGACGAGTTGACGTTCATGGT  
 L H H V S S L G E I F N G L L N C K Y H -

CAAATGCCCTTACATAATTGGCACAAATGAACCAGGCTACCTCTTCGACTCCCCAACCTGCA  
 6421 -----+-----+-----+-----+-----+-----+-----+ 6480  
 GTTTACGGAATGTATTAACCGTGTACTGGTCCGATGGAGAAGCTGAGGGTTGGACGT  
 K C P Y I I G T M N Q A T S S T P N L Q -

GCTTCACCATAACTTCAGATGGGTGCTTGCCAAACACACGGAGCCTGTGAAGGGTTT  
 6481 -----+-----+-----+-----+-----+-----+-----+ 6540  
 CGAAGTGGTATTGAAGTCTACCCACGAAACACGGTTGGTGTGCCTCGGACACTTCCAAA  
 L H H N F R W V L C A N H T E P V K G F -

CCTTGGCCGATTCTGAGGAGGAAGCTCATGAAACAGAGATCAGTGGCGGGTGGCAA  
 6541 -----+-----+-----+-----+-----+-----+-----+ 6600  
 GGAACCGGCTAAGGACTCCTCCTCGAGTACCTTGTCTTAGTCACCCGCCACCGCGTT  
 L G R F L R R K L M E T E I S G R V R N -

TATGGAGCTGGTAAAATCATTGACTGGATTCCCAAGGTCTGGCATCACCTCAACCGCTT  
 6601 -----+-----+-----+-----+-----+-----+-----+ 6660  
 ATACCTCGACCATTAGTAAGTACTGACCTAACGGTTCCAGACCGTAGTGGAGTTGGCGAA  
 M E L V K I I D W I P K V W H H L N R F -

CCTGGAGGCTCACAGTTCCTCGGACGTACCATCGGCCCCGGCTTCTGTGCATGCC  
 6661 -----+-----+-----+-----+-----+-----+-----+ 6720  
 L E A H S S S D V T I G P R L F L S C P -

CATCGATGTGGACGGCTCGAGAGTGTGGTTACCGACTTGTGGAACATTCCATTATCCC  
 6721 -----+-----+-----+-----+-----+-----+-----+ 6780  
 I D V D G S R V W F T D L W N Y S I I P -

CTATCTCCTGGAAGCCGTCAAGAGAAGGACTCCAGCTCTATGGAAGGCGGCCCTGGGA  
 6781 -----+-----+-----+-----+-----+-----+-----+ 6840  
 Y L L E A V R E G L Q L Y G R R A P W E -

GGATCCTGCCAAGTGGGTGATGGACACATATCCATGGCAGCCACAAACAGCACGA  
 6841 -----+-----+-----+-----+-----+-----+-----+ 6900  
 P P A K W V M D T Y P W A A S P Q Q H E -

GTGGCCTCCCTGCTGCAGTTACGGCCTGAGGATGTCGGCTTCGACGGCTACTCCATGCC  
 6901 -----+-----+-----+-----+-----+-----+-----+ 6960  
 W P P L L Q L R P E D V G F D G Y S M P -

TCGGGAGGGATCGACAAGCAAGCAGATGCCCGAGTGTGCTGAAGGTGACCCGCTGAT  
 6961 -----+-----+-----+-----+-----+-----+-----+ 7020  
 R E G S T S K Q M P P S D A E G D P L M -

GAACATGCTGATGAGGCTGCAGGAGGCAGCCAACACTCCAGCCCCAGAGCTATGACAG  
 7021 -----+-----+-----+-----+-----+-----+-----+ 7080  
 N M L M R L Q E A A N Y S S P Q S Y D S -

Fig. 17 (cont'd 7)

09/914549

94/124

CGACTCCAACAGCAACAGCCATCACGATGACATCTGGACTCCTCTTGGAGTCCACTCT  
7081 -----+-----+-----+-----+-----+-----+-----+-----+ 7140  
D S N S N H H D D I L D S S L E S T L -  
  
GTGACAGGGGCCCGGAGCCCAGCGCCCTCCTCTTCCTCACCGCATTCCACCTGCATCC  
7141 -----+-----+-----+-----+-----+-----+-----+-----+ 7200  
\*  
  
CCCACATCACCCCTGAAGATGACTTCCTGAGCCAGCCCCAGCCACAGCCTTAGAGCTGCG  
7201 -----+-----+-----+-----+-----+-----+-----+-----+ 7260  
  
GGAACACCGAGACCCCCCGTCCTTCAGCCTCGACCTGGTGAGGCATCCGGCCAGCT  
7261 -----+-----+-----+-----+-----+-----+-----+-----+ 7320  
  
GCCTGCGGACCGCTTCCTCACAGCGAGAACTGCACTACCTCTGTTGTACTTTAATTA  
7321 -----+-----+-----+-----+-----+-----+-----+-----+ 7380  
  
TTGTTTGCCTGTTGCTGTGACCTCCCTAACGACACTGAAGATACTTCCTGGAAAGGAT  
7381 -----+-----+-----+-----+-----+-----+-----+-----+ 7440  
  
CATCGCCGTTGAAATGAAAAGAGAGACAGAGAGAGAAAAAAAAGAGAACCCACATGAA  
7441 -----+-----+-----+-----+-----+-----+-----+-----+ 7500  
  
GCTCTGAAACCAAACAGCATTGCCATGAGCTTCCCAGAGACAGAACAGAGACTGGAGCAA  
7501 -----+-----+-----+-----+-----+-----+-----+-----+ 7560  
  
AGTCGGAAACACAGAGAACGGCACGGCTCCCTCAGCACAGACCCCTCCAGACTGGTCTCA  
7561 -----+-----+-----+-----+-----+-----+-----+-----+ 7620  
  
GAGCCGTGCCACCCACCCTCCACACAGCCGGCACAGGGAGAACTGGTGCTAACCAGGG  
7621 -----+-----+-----+-----+-----+-----+-----+-----+ 7680  
  
TGCTTGCCTTGGTCACGTTAACGCACTACAGAGCTACGACACAGGGAACCTTAGGAGC  
7681 -----+-----+-----+-----+-----+-----+-----+-----+ 7740  
  
AAATAAACCGTGTTCATGTTTTAAAAAAAAAAAAAAA  
7741 -----+-----+-----+-----+-----+-----+-----+ 7783

Fig. 17 (cont'd 8)

**T3-cDNA sequence and T3 protein encoded  
therein (isoform 2)**

1	AGCAGGGAGAGGGAGGGAGTGTGCCGTCTTCCTGCAAGGGCAGTGCCCCAGCCTCAGC	60
S R E R G G S V P S L L Q G Q C P S L S -		
61	CACACTTCTGATCTGCAGTCCAACAGACCTTCTAGCATGCCAAAGAGAACCTGGGGTG	120
H T S D L Q S N R P F * H A K E N L G V -		
121	CCAGGGGTCTCAGAGCTCACACTGCACCTGTGGCACCCACAGCGAGTAGCCATCCGTG	180
P G G P Q S S H C T C G T H S E * P S V -		
181	AGCCGAGGAAACTGTACACAGATCTACACAGACTGGCCAATCATTACCTAGCCAAATCC	240
S R G N C T Q I Y T D W A N H Y L A K S -		
241	GGCCACAAGCGTCTCATCAAGGATCTCAGCAAGATGTGACAGATGGCGTCCTCGGCC	300
G H K R L I K D L Q Q D V T D G V L L A -		
301	CAGATTATCCAGGTTGTGGCAAATGAAAAGATTGAAGACATCAATGGCTGTCCGAAGAAC	360
Q I I Q V V A N E K I E D I N G C P K N -		
361	AGATCCCAAATGATTGAAAACATAGATGCCTGTTGAATTCTGGCAGCTAAGGAAATA	420
R S Q M I E N I D A C L N F L A A K G I -		
421	AACATCCAGGGCTGTCTGCAGAAGAGATCAGGAATGGAAACCTCAAGGCCATTCTAGGC	480
N I Q G L S A E E I R N G N L K A I L G -		
481	CTCTTCTTCAGCCTCTCCGATACAAGCAGCAGCAGCAGCAGCCCCAGAAGCAGCACCTC	540
L F F S L S R Y K Q Q Q Q P Q K Q H L -		
541	TCCTCACCTCTGCCGCCGCCGTATCCCAGGTGCCGGGGCCCCCTCCCAGTGCAGGCT	600
S S P L P P A V S Q V A G A P S Q C Q A -		
601	GGCACCCCTCAGCAGCAGGTGCCAGTCACTCCCCAAGCCCCGTGCCAGCCTCACCAAGCCA	660
G T P Q Q V P V T P Q A P C Q P H Q P -		

Fig. 18

96/124  
T3 murine cDNA

ATGAGAAGAGCCGAACAATGAGTCGGTCAGGCTCCTCCGGATGGTTGAGGAAGTTC  
 1 E K S R T M S R S G S F R D G F E E V H - 60

ATGGATCCTCCCTGTCCTGGTTCCAGCACATCCTCCATCTACTCCACGCCAGAAAGAAA  
 61 G S S L S L V S S T S S I Y S T P E E K - 120

AATGCCAGTCAGAGATTGAAAGCTGAGGCAGAACCTGGATGCCTCCCAGGAAAGGTGT  
 121 C Q S E I R K L R R E L D A S Q E K V S - 180

AATGCCAGTCAGAGATTGAAAGCTGAGGCAGAACCTGGATGCCTCCCAGGAAAGGTGT  
 121 C Q S E I R K L R R D V D A S Q E K V S - 180

CTGCGCTGACTACCCAGCTGACTGCAAATGCTCACCTGTGGCAGCCTCGAGCAGAGTC  
 181 A L T T Q L T A N A H L V A A F E Q S L - 240

TGGGAAACATGACCACAGGCTACAGAGTTAACTATGACCGCTGAGCAGAAGGATTCA  
 241 G N M T I R L Q S L T M T A E Q K D S E - 300

AACTGAACGAGTTAAGAAAAACCACGAGCTGCTGAAGAACAGAACATGCAGCTGCCAGG  
 301 L N E L R K T I E L L K K Q N A A A Q A - 360

CTGCCATTAATGGAGTGATTAACACGCCAGAGCTCAACTGCAAAGGAAATGGCAGTGCA  
 361 A I N G V I N T P E L N C K G N G S A R - 420

GGCTACAGACCTACGCATCCGAGCAACACTCCTCCGACAGTGTCTCCAGTATCAATAGC  
 421 L Q T Y A S A A T L L R Q C L Q Y Q \* R - 480

GCCACCAGCCACTCAAGTGTG  
 481 H Q P L K C - 501

Fig. 19

T2

CAGGCTCTCCAACCGCTCGTA  
S L S N R S SAGCCTCTCCAACCGCTCGTAC  
S L S N R S YGECCTCTCCAACCGCTCGTACC  
S L S N R S YCCTCTCCAACCGCTCGTACCC  
L S N R S YCTCTCCAACCGCTCGTACCCCT  
L S N R S Y PCTCCAACCGCTCGTACCCCT  
S N R S Y PTCCAACCGCTCGTACCCCTCTG  
S N R S Y P LCCAACCGCTCGTACCCCTCTGT  
S N R S Y P LCAACCGCTCGTACCCCTCTGTC  
N R S Y P LAACCGCTCGTACCCCTCTGTCA  
N R S Y P L SACCGCTCGTACCCCTCTGTCAT  
N R S Y P L SCCGCTCGTACCCCTCTGTCATG  
R S Y P L SCGCTCGTACCCCTCTGTATGG  
R S Y P L S WGCTCGTACCCCTCTGTATGGC  
S Y P L S WCTCGTACCCCTCTGTATGGCG  
S Y P L S WTCGTACCCCTCTGTATGGCGC  
S Y P L S W RCGTACCCCTCTGTATGGCGCT  
S Y P L S W RGTACCCCTCTGTATGGCGCTA  
Y P L S W RTACCCCTCTGTATGGCGCTAT  
Y P L S W R YACCCCTCTGTATGGCGCTATG  
Y P L S W R Y

T2

97/124

CCTCCTCCACCTACTCCTCAC  
A S S T Y S SCTCCTCCACCTACTCCTCAC  
A S S T Y S STCCTCCACCTACTCCTCACAA  
A S S T Y S S QCTCCACCTACTCCTCACAAAT  
A S S T Y S S QCTCCACCTACTCCTCACAAATC  
A S T Y S S Q ICCACCTACTCCTCACAAATCC  
A S T Y S S Q ICACCTACTCCTCACAAATCCG  
A T Y S S Q IACCTACTCCTCACAAATCCGG  
A T Y S S Q I RCCTACTCCTCACAAATCCGGA  
A T Y S S Q I RCTACTCCTCACAAATCCGGA  
A Y S S Q I RTACTCCTCACAAATCCGGAG  
A Y S S Q I R KACTCCTCACAAATCCGGAAGC  
A Y S S Q I R KCTCCTCACAAATCCGGAAGCT  
A S S Q I R KTCCTCACAAATCCGGAAGCTT  
A S S Q I R K LCCTCACAAATCCGGAAGCTTC  
A S S Q I R K LCTCACAAATCCGGAAGCTTC  
A S Q I R K LTCACAAATCCGGAAGCTTCGT  
A S Q I R K L RCACAAATCCGGAAGCTTCGTA  
A S Q I R K L RACAAATCCGGAAGCTTCGTAG  
A Q I R K L RT2  
09/914549  
AGAAGAAAAAAAAGAGTTGGC  
K K K K S WGAAGAAAAAAAAGAGTTGGC  
K K K K S WAAGAAAAAAAAGAGTTGGC  
K K K K S W LAGAAAAAAAAGAGTTGGCTTC  
K K K K S W LGAAAAAAAAGAGTTGGCTTCG  
K K K S W LAAGAAAAAGAGTTGGCTTCGA  
K K K S W L RAAGAAAAAGAGTTGGCTTCGA  
K K K S W L RAAGAAAAAGAGTTGGCTTCGA  
K K S W L RAAGAGAGTTGGCTTCGAAGT  
K K S W L R SAAGAGAGTTGGCTTCGAAGT  
K K S W L R SAAGAGAGTTGGCTTCGAAGT  
K S W L R SAAGAGAGTTGGCTTCGAAGT  
K S W L R S SAGAGTTGGCTTCGAAGTTCC  
K S W L R S SGAGTTGGCTTCGAAGTTCC  
S W L R S SAGTTGGCTTCGAAGTTCC  
S W L R S S FGTTGGCTTCGAAGTTCC  
S W L R S S FTTGGCTTCGAAGTTCC  
W L R S S FTGGCTTCGAAGTTCC  
W L R S S F NGGCTTCGAAGTTCC  
W L R S S F NGCTTCGAAGTTCC  
L R S S F N

09/914549

T2

T2 98/124

CTCCATCAAGTCCTCCACCTC S I K S S T	AGTTGGAGGTGGACCTGCTGG L E V D L L	ATGACACCCAACCATGGCTTT M T P N H G F
TCCATCAAGTCCTCCACCTCG S I K S S T S	GTTGGAGGTGGACCTGCTGG L E V D L L	TGACACCCAACCATGGCTTT M T P N H G F
CCATCAAGTCCTCCACCTCGT S I K S S T S	TTGGAGGTGGACCTGCTGG L E V D L L E	GACACCCAACCATGGCTTT T P N H G F
CATCAAGTCCTCCACCTCGTC I K S S T S	TGGAGGTGGACCTGCTGG L E V D L L E	ACACCCAACCATGGCTTT T P N H G F H
ATCAAGTCCTCCACCTCGTC I K S S T S S	GGAGGTGGACCTGCTGG E V D L L E	CACCCAACCATGGCTTT T P N H G F H
TCAAGTCCTCCACCTCGTCT I K S S T S S	GAGGTGGACCTGCTGG E V D L L E A	ACCCAACCATGGCTTT P N H G F H
CAAGTCCTCCACCTCGTCTC K S S T S S	AGGTGGACCTGCTGG E V D L L E A	CCCCAACCATGGCTTT P N H G F H L
AAGTCCTCCACCTCGTCTC K S S T S S S	GGTGGACCTGCTGG V D L L E A	CACCATGGCTTT P N H G F H L
AGTCTCCACCTCGTCTC K S S T S S S	GTGGACCTGCTGG V D L L E A E	CAACCATGGCTTT N H G F H L
GTCCTCCACCTCGTCTCCG S S T S S S	TGGACCTGCTGG V D L L E A E	AACCATGGCTTT N H G F H L S
TCTCCACCTCGTCTCCG S S T S S S V	GGACCTGCTGG D L L E A E	ACCATGGCTTT N H G F H L S
CCTCCACCTCGTCTCCGTGG S S T S S S V	GACCTGCTGG D L L E A E N	CCATGGCTTT H G F H L S
CTCCACCTCGTCTCCGTGGGG S T S S S V	ACCTGCTGG D L L E A E N	CATGGCTTT H G F H L S F
TCCACCTCGTCTCCGTGGGG S T S S S V G	CCTGCTGG L L E A E N	ATGGCTTT H G F H L S F
CCACCTCGTCTCCGTGGGGCA T S S S V G	CTGCTGG L L E A E N D	TGGCTTT G F H L S F
CACCTCGTCTCCGTGGGCAC T S S S V G	TGCTGG L L E A E N D	GGCTTT G F H L S F R
ACCTCGTCTCCGTGGGCAC T S S S V G G	GCTGG L E A E N D	GCTTT G F H L S F R
CCTCGTCTCCGTGGGCAC T S S S V G G	CTGGA L E A E N D R	CTTT F H L S F R
CTCGTCTCCGTGGGCAC S S S V G G	TGGA L E A E N D R	TTT F H L S F R M
TCGTCTCCGTGGGCAC S S S V G G T	GGA E A E N D R	TTCA F H L S F R M
CGTCCTCCGTGGGCAC S S V G G T	GAAG E A E N D R L	TCAC H L S F R M

Fig. 20 (cont'd 1)

09/914549

T2

TAAGGTAAAAATGAAAAAC  
AAAGGTAAAATGAAAAACA  
AAAGGTAAAATGAAAAACAA  
AAGGTAAAATGAAAAACAAA  
ACGTAAAATGAAAAACAAA  
GGTAAAATGAAAAACAAAAA  
GTAAAATGAAAAACAAAAC  
TAAAATGAAAAACAAAACA  
AAAAATGAAAAACAAAACA  
AAAATGAAAAACAAAACAAG  
AAATGAAAAACAAAACAAGC  
AATGAAAAACAAAACAAGCA  
ATGAAAAACAAAACAAGCAA  
TGAAAAACAAAACAAGCAA  
GAAAACAAAACAAGCAAAC  
AAAAACAAAACAAGCAAACA  
AAAACAAAACAAGCAAACAA  
AAACAAAACAAGCAAACAA  
AACAAAACAAGCAAACAAAC  
ACAAAACAAGCAAACAAACA

99/124

T2

CTCTAATTAAACATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
CCTCTAATTAAACATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
CTCTAATTAAACATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
TCTAATTAAACATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
CTAATTAAACATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
TAATTAAACATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
AATTAACATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
ATTAACATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
TTAACATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
TTAACATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
TAACAATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
AACATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
ACAATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
CAATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
AATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
ATGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
TCGCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
GCATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
CATAGAGTCATAAAACCTACTTTTTAAAAAAA  
ATAAGAGTCATAAAACCTACTTTTTAAAAAAA  
TAAGAGTCATAAAACCTACTTTTTAAAAAAA

Fig. 20 (cont'd 2)

09/914549

100/124

T3

ACTGGGCAATCATTACCTAG  
W A N H Y L

CTGGGCAATCATTACCTAGC  
W A N H Y L

TGGGCAATCATTACCTAGCC  
W A N H Y L A

GGGCAATCATTACCTAGCCA  
W A N H Y L A

GGCCAATCATTACCTAGCCAA  
A N H Y L A

GCCAATCATTACCTAGCCAAA  
A N H Y L A K

CCAATCATTACCTAGCCAAAT  
A N H Y L A K

CAATCATTACCTAGCCAAATC  
N H Y L A K

AATCATTACCTAGCCAAATCC  
N H Y L A K S

ATCATTACCTAGCCAAATCCG  
N H Y L A K S

TCATTACCTAGCCAAATCCGG  
H Y L A K S

CATTACCTAGCCAAATCCGGC  
H Y L A K S G

ATTACCTAGCCAAATCCGGCC  
H Y L A K S G

TTACCTAGCCAAATCCGGCCA  
Y L A K S G

TACCTAGCCAAATCCGGCCAC  
Y L A K S G H

ACCTAGCCAAATCCGGCCACA  
Y L A K S G H

CCTAGCCAAATCCGGCCACAA  
L A K S G H

CTAGCCAAATCCGGCCACAAG  
L A K S G H K

TAGCCAAATCCGGCCACAAGC  
L A K S G H K

AGCCAAATCCGGCCACAAGCG  
A K S G H K

GCCAAATCCGGCCACAAGCGT  
A K S G H K R

T3

CGGCCACAAGCGTCTCATCAG  
G H K R L I

GGCCACAAGCGTCTCATCAGG  
G H K R L I R

GCCACAAGCGTCTCATCAGGG  
G H K R L I R

CCACAAGCGTCTCATCAGGGGA  
H K R L I R

CACAAGCGTCTCATCAGGGAT  
H K R L I R D

ACAAGCGTCTCATCAGGGATC  
H K R L I R D

CAAGCGTCTCATCAGGGATCT  
K R L I R D

AAGCGTCTCATCAGGGATCTC  
K R L I R D L

AGCGTCTCATCAGGGATCTCC  
K R L I R D L

GCGTCTCATCAGGGATCTCCA  
R L I R D L

CGTCTCATCAGGGATCTCCAG  
R L I R D L Q

GTCTCATCAGGGATCTCCAGC  
R L I R D L Q

TCTCATCAGGGATCTCCAGCA  
L I R D L Q

CTCATCAGGGATCTCCAGCAA  
L I R D L Q Q

TCATCAGGGATCTCCAGCAAG  
L I R D L Q Q

CATCAGGGATCTCCAGCAAGA  
I R D L Q Q

ATCAGGGATCTCCAGCAAGAT  
I R D L Q Q D

TCAGGGATCTCCAGCAAGATG  
I R D L Q Q D

CAGGGATCTCCAGCAAGATGT  
R D L Q Q D

AGGGATCTCCAGCAAGATGTG  
R D L Q Q D V

GGGATCTCCAGCAAGATGTGA  
R D L Q Q D V

T3

CTGAAATGCAGTCCAGACTTC  
E M Q S R L

TGAAATGCAGTCCAGACTTCC  
E M Q S R L

GAAATGCAGTCCAGACTTCCA  
E M Q S R L P

AAATGCAGTCCAGACTTCCAG  
E M Q S R L P

AATGCAGTCCAGACTTCCAGG  
M Q S R L P

ATGCAGTCCAGACTTCCAGGT  
M Q S R L P G

TGCAGTCCAGACTTCCAGGT  
M Q S R L P G

GCAGTCCAGACTTCCAGGTCC  
Q S R L P G

CAGTCCAGACTTCCAGGTCT  
Q S R L P G P

AGTCCAGACTTCCAGGTCTTA  
Q S R L P G P

GTCCAGACTTCCAGGTCTTAC  
S R L P G P

TCCAGACTTCCAGGTCTTAC  
S R L P G P T

CCAGACTTCCAGGTCTTACCG  
S R L P G P T

CAGACTTCCAGGTCTTACCGC  
R L P G P T

AGACTTCCAGGTCTTACCGC  
R L P G P T A

GACTTCCAGGTCTTACCGC  
R L P G P T A

ACTTCCAGGTCTTACCGC  
L P G P T A

CTTCCAGGTCTTACCGC  
L P G P T A R

TTCCAGGTCTTACCGC  
L P G P T A R

TCCAGGTCTTACCGC  
P G P T A R

CCAGGTCTTACCGC  
P G P T A R V

Fig. 20 (cont'd 3)

09/914549

T3 101/124

CGGGGCACTAGTGTCTGCCAC  
R G S S V C H

GGGCAGTAGTGTCTGCCACCG  
R G S S V C H

GGGCAGTAGTGTCTGCCACGT  
G S S V C H

GGCAGTAGTGTCTGCCACGTG  
G S S V C H V

GCAGTAGTGTCTGCCACGTGG  
G S S V C H V

CAGTAGTGTCTGCCACGTGGAA  
S S V C H V

AGTAGTGTCTGCCACGTGGAC  
S S V C H V D

GTAGTAGTCTGCCACGTGGACG  
S S V C H V D

AGTAGTGTCTGCCACGTGGACGT  
S V C H V D

GTGTCTGCCACGTGGACGTCT  
S V C H V D V

TGTCTGCCACGTGGACGTCTC  
V C H V D V V

GTCTGCCACGTGGACGTCTCA  
V C H V D V S

TCTGCCACGTGGACGTCTCAG  
V C H V D V S

CTGCCACGTGGACGTCTCAGA  
C H V D V S

TGCCACGTGGACGTCTCAGAC  
C H V D V S D

GCCACGTGGACGTCTCAGACA  
C H V D V S D

CCACGTGGACGTCTCAGACAA  
H V D V S D

CACGTGGACGTCTCAGACAAG  
H V D V S D K

ACGTGGACGTCTCAGACAAGG  
H V D V S D K

CGTGGACGTCTCAGACAAGG  
V D V S D K

AGAACAGAGTGGTTCCGCCA  
K Q S G S A

GAAGCAGAGTGGTTCCGCCAC  
K Q S G S A

AAGCAGAGTGGTTCCGCCACC  
K Q S G S A T

AGCAGAGTGGTTCCGCCACCG  
K Q S G S A T

GCAGAGTGGTTCCGCCACCGG  
Q S G S A T

CAGAGTGGTTCCGCCACCGGC  
Q S G S A T G

AGAGTGGTTCCGCCACCGGCC  
Q S G S A T G

GAGTAGTGGTTCCGCCACCGGCC  
S G S A T G

AGTAGTGGTTCCGCCACCGGCC  
S G S A T G L

GTGGTTCCGCCACCGGCC  
S G S A T G L

TGGTTCCGCCACCGGCC  
G S A T G L

GGTTCCGCCACCGGCC  
G S A T G L A

GTTCCGCCACCGGCC  
G S A T G L A

TTCCGCCACCGGCC  
S A T G L A

TCCGCCACCGGCC  
S A T G L A M

CCGCCACCGGCC  
S A T G L A M

CGCCACCGGCC  
A T G L A M

GCCACCGGCC  
A T G L A M I

CCACCGGCC  
T G L A M I

CACCGGCC  
T G L A M I

ACCGGCC  
T G L A M I T

Fig. 20 (cont'd 4)

T3

T3 102/124

09/914549

GGTCTGGTCAACCAAACAGAC  
G L V N Q T D

GTCCTGGTCAACCAAACAGACA  
G L V N Q T D

TCTGGTCAACCAAACAGACAA  
L V N Q T D

CTGGTCAACCAAACAGACAAG  
L V N Q T D K

TGGTCAACCAAACAGACAAGG  
L V N Q T D K

GGTCAACCAAACAGACAAGGA  
V N Q T D K

GTGAAACCAAACAGACAAGGAG  
V N Q T D K E

TCAACCAAACAGACAAGGAGA  
V N Q T D K E

CAACCAAACAGACAAGGAGAA  
N Q T D K E

CAACCAAACAGACAAGGAGAAA  
N Q T D K E K

ACCAAACAGACAAGGAGAAAAG  
N Q T D K E K

CCAAACAGACAAGGAGAAAAGG  
Q T D K E K

CAAACAGACAAGGAGAAAAGGC  
Q T D K E K G

AAACAGACAAGGAGAAAAGGC  
Q T D K E K G

AACAGACAAGGAGAAAAGGCAT  
T D K E K G

ACAGACAAGGAGAAAAGGCATC  
T D K E K G I

CAGACAAGGAGAAAAGGCATCT  
T D K E K G I

ACAAGGAGAAAAGGCATCTC  
D K E K G I

GACAAGGAGAAAAGGCATCTCA  
D K E K G I S

ACAAGGAGAAAAGGCATCTCAT  
D K E K G I S

CAAGGAGAAAAGGCATCTCATC  
K E K G I S

TTCATGGATCCTCACTCTCCT  
H G S S L S

TCATGGATCCTCACTCTCCTT  
H G S S L S

CATGGATCCTCACTCTCCTTGG  
H G S S L S L

ATGGATCCTCACTCTCCTTGG  
H G S S L S L

TGGATCCTCACTCTCCTTGG  
G S S L S L

GGATCCTCACTCTCCTTGG  
G S S L S L V

GATCCTCACTCTCCTTGG  
G S S L S L V

ATCCTCACTCTCCTTGG  
S S L S L V

TCCTCACTCTCCTTGG  
S S L S L V S

CCTCACTCTCCTTGG  
S S L S L V S

CTCACTCTCCTTGG  
S L S L V S

TCACTCTCCTTGG  
S L S L V S S

CACTCTCCTTGG  
S L S L V S S

ACTCTCCTTGG  
L S L V S S

CTCTCCTTGG  
L S L V S S T

TCTCCTTGG  
L S L V S S T

CTCCTTGG  
S L V S S T

TCCTTGG  
S L V S S T S

CCTTGG  
S L V S S T S

CTTGG  
L V S S T S

TTGG  
L V S S T S S

CTCCTTGGTTCCAGCACATC  
S L V S S T

TCCTTGGTTCCAGCACATCG  
S L V S S T S

CCTGGTTTCCAGCACATCGT  
S L V S S T S

CTTGGTTTCCAGCACATCGTC  
L V S S T S

TTGGTTTCCAGCACATCGTC  
L V S S T S S

TGTTTCCAGCACATCGTCAG  
L V S S T S S

GGTTTCCAGCACATCGTCAG  
V S S T S S

GTTTCCAGCACATCGTCAG  
V S S T S S V

TTTCCAGCACATCGTCAG  
V S S T S S V

TTCCAGCACATCGTCAG  
S S T S S V

TCCAGCACATCGTCAG  
S S T S S V Y

CCAGCACATCGTCAG  
S S T S S V Y

CAGCACATCGTCAG  
S T S S V Y

AGCACATCGTCAG  
S T S S V Y S

GCACATCGTCAG  
S T S S V Y S

CACATCGTCAG  
T S S V Y S

ACATCGTCAG  
T S S V Y S T

CATCGTCAG  
T S S V Y S T

ATCGTCAG  
S S V Y S T

TCGTCAG  
S S V Y S T P

CGTCAG  
S S V Y S T P

T3

GGAGAGAACTGGGTCAATGAGTTACGCAGCTCC  
K N W V N E L R S S

Fig. 20 (cont'd 5)

09/914549

AAGAAGAAGAGGAAGAACTGGGTCAATGAGTTACGCAGCTCCTCAAG  
K K K R K N W V N E L R S S F K

103/124

T3

TCTCTAATTCAAGAACATGGATA

T3

AGGAGATGAAGCTGACAGATATCCGCTTAGAAGCTCT

T3

GATTCCAGACCACACGTCTTCTTATCG

Fig. 20 (cont'd 6)

## Alignment of the T protein family

Note: The N-terminus of protein T2 was omitted in the alignment, since it has no significant homology to the T protein and the T3 protein.

T	---MDLSSEMNRHGKNPVSHKLEDQK-----	KIYTDWANHYLAKSGHKRLIKDLQ	60
T3	NQPERLNSQVLQLQE PAGEGLPLRKSGSVENGFDTQI	YTDWANHYLAKSGHKRLIRDLQ	
T2	-----	*****	
	* * . . . * * . *	*****	
T	QDIADGVLLAEIIQIIANEKVEDINGCPRSQSQMIENVDVCLSFLAARGVNQGLSAEEI	120	
T3	QDVTDGVLLAQIIQVVANEKIEDINGCPKNRSQMIENIDACLNFLAAKGINIQGLSAEEI		
T2	-----	*****	
	* * . ***** . *** . **** . ***** .	***** . * * * * . * . * . *****	
T	RNGNLKAILGLFFSLSRYKQQQ-HHQQQYYQS-----	LVELQQRVT	180
T3	RNGNLKAILGLFFSLSRYKQQQQQPQKQHLSSPLPPAVSQAGAPSQCQAGTPQQQVPVT		
T2	-----	***** . * . * * .	
	*****	* *	
T	HASP-----PSEASQAKTQQDMQSRLPGP-SRVPAAGSSSKVQGASNLN--RRSQSFNSI	240	
T3	PQAPCQPHQPAHQSKAQAEMQSRLPGPTARVSAAGSEAKTRGGTTANNRRSQSFNNY		
T2	-----	***** . * . * . * . ***** . * * . * . * . . * . *****	
T	DKNK----PP-----	300	
T3	DKSKPVTSPPPPSSHEKEPLASSASSHPGMSDNAPASLESGSSSTPTNCSTSSAIQPG		
T2	-----	** *	
	**		
T	-----N-----	360	
T3	AATKPWRSKSLSVKHSATVMSLVKPPGPEAPRPTPEAMKPAPNNQKSMLEKLKLFNSKG		
T2	-----	*	
T	-----	420	
T3	GSKAGEGPGRDTSCERLETLPSEEELAASRMLTTVGPASSSPKIALKGIAQRTFS		
T2	-----		
T	--YAN-----GNEK-----	480	
T3	RALTNKKSSLKGNEKEKEQQREKDKEKSNDLAKRASVTERLDLKEEPKEDPSGAAPMEM		
T2	-----	.* ***	
T	-----	540	
T3	PKKSSKIASFIPKGGKLNSAKKEPMAPSHSGIPKPGMKSMPGKSPSAPAPSKEGERSRSG		
T2	-----		
T	-----	600	
T3	KLSSGLPQQKPQLDGRHSSSSSLASSEGKGPGGTTLNHSISSQTVGSGVTTQTTGSNT		
T2	-----		

Fig. 21

T -----GEDPETRRMRTVKNIADLRQNLEETMSSLRGQTQISHSTLETTFDSTVTTEVNGRTIP 720  
T3 VSVQLPQPQQQYNHPNTATVAPFLYRSQTDTEGNVTAESSSTGVSVEPSHFTKTGQPALE  
T2 -----  
  
T ---ELTGEDPEARRLRTVKNIADLRQNLEETMSSLRGQTQVTHSTLETTFDTNVTTEMGRSIL 720  
T3 -----DPESQRKRTVQNVLSDLRQNLEETMSSLRGSGVTHSSLEMTCYD--SDDANPRSVS  
T2 -----  
\*\*\*\*\* . \* \* . \*  
  
T NLTSRPTPMTWRLGQACPRLQAGDAPS LGAGYP-RSGTSRFIHTDPSRFMYTTPLRRAAV 780  
T3 SLTGRPTPLSWRLGQSSPRLQAGDAPS MGNGYPPRANASRFINTESGRYVYSAPLRRQLA  
T2 SLSNRSYPLSWRYGQSSPRLQAGDAPS VGGSCRSEGT PAWYMHGERAHYSHTMPMR--SP  
\* . \* \* . \*  
  
T SRLGNMSQIDMSEKA-SSLDLMS-SEVDVGGYMSDGDILGKSLRTDDINSGYMTDGGLNL 840  
T3 SRGSSVCHVDVSDKA-GDEMDLEGISMDAPGYMSDGDVLSKNIRTDDITSGYMTDGGLGL  
T2 SKLSHISRLELVESDLSDDEVDLK-----SGYMSDSDLMGKTMTEDDD----ITTG----  
\* . . . . . \* . \* \* \* \* . \* . \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \*  
  
T YTRSLNRIPD-TATS RDIIQRGVHDVTVDADSWDDSSVSSGLSDTL DNI STDDLNTTSS 900  
T3 YTRRLNRLPDGMMAVVRETLQRNTSLGLGDADSWDDSSVSSGISDTIDNLSTDDINTSSS  
T2 -----WDESSSISSGLSDASDNLSSEEFNASSS  
\*  
  
T VSSYSNITVPSRKN--TQLRTDSEKRSTTDET--WDSP--EELKKPE--EDFD SHGDAG- 960  
T3 ISSYANTPASSRKN--LDVQ TD A E KHS QVERNSLWSG--DDVKKSDGGSDSGIKMEPG-  
T2 LNSLPSTPTASRRNSTIVLRTDSEKRSLAE S GLS WFSESEEKAPK KLEYDGS LKMEPGT  
\* . \* \* . \* . \* \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \*  
  
T GKWKT VSSGLPEDPEK-AGQKASLSV S QTGS WRRGMSAQGGAPS--RQKAGTSALKTP- 1020  
T3 SKWRRNPSDVSDES DKSTSGKKNPVI S QTGS WRRG MTAQVGITM PRTKASA PAGALKTPG  
T2 SKWRRERPESCDDSSKG GELKKPISLGH PGSLKKGKTPPVAVTSP--IHTAQSALKVAG  
\* . . . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \*  
  
T -GKTDDAKASEKGKAPLK GSSLQRSPSDAGKSSGDEGKK--PPSGIGRSTATSSFGFKKP 1080  
T3 TGKTDDAKVSEKG RLS PKASQVKRSPSDAGRSSGDESKKPLPSSSRTPTANANSFGFKKP  
T2 ---KPEGKATDKGKLAVKNTGLQRSS SDAGRDR L SDAKK--PPSGIARPSTSGSF GYKKP  
\* . . \* . \* . . . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \*  
  
T SG-VGSSAMITSSGATITSGSATLGKIPKSAAIGGKS NAGR KTS LDGSQ NQDDV VLHVSS 1140  
T3 SGSATGLAMITASGVTVTSRSATLGKIPKSSALVSRS-AGR KSSMDGAQ NQDDG YLA LSS  
T2 PP-ATGTATVMQTG----GSATLSKIQKSSGIPVKPVNGRKTS LDVSNSAEPGFLAPGA  
\* . . \* . \* . \* \* \* \* \* . . . \* \* . \* . \* . \* . \* . \* . \* . \* . \*  
  
T KTTLQYRSLPRPSKSSTSGIPGR-GGHR SSTSSID-SNVSSKSAGATT SKLREPTKIGSG 1200  
T3 RTNLQYRSLPRPSKSNSR--NG--AGNRSSTSSID-SNISSKSAGLPVPKLREPSK TALG  
T2 RSNIQYRSLPRPAKSSSMSVTGGRRGGPRPVSS IDPSLLSTKQGGLTPS RLKEPTKVASG  
\* . \* \* \* \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \*  
  
T RSSPVTVNQTDKEKEKVAVSDSESVLSG-SPKSSPT SASACG-AQGLRQPGSKYPDIA S 1260  
T3 SSLPGLVNQTDKEKG--ISSDNE SVASCN-SVKVNPA AQP VSSPAQ TSLQPGAKY PDVAS  
T2 RTTPAPVNQTDREKE---KAKAKAVAL DSDNISLKSIGSPESTPKNQASHPTATKLAELP  
\* . \* \* \* \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \*  
  
T PTF RRLFGAKAGGKSASAPNTEGVKSSVMPSPSTTLARQGSLES PSSGTGSMGSAGGLS 1320  
T3 PTL RRLFGGKP-TKQVPIATAENMKNSVVISNP HATMTQQGNLDSPS-GSGVLS--S  
T2 PTPLRAT-AKSFVKKPPSLANLDKVN-SNSLDLPSSSDTTHASKV PDLHATSSAS-----

```

T  GSSSPLFNKPSDLTTDVISLSHSLASSPASVHSFTSGGLVWAANMSSSSAGSKDTPSYQS 1380
T3 GSSSPLYSKNVDLN-----QSPLOSSPSSAHSAPSNSLTWGTNASSSSAVSKDGLGFQS
T2 --GGPLPS-----CFTPSPAPILNINSASFSQLMLEMSGFSVPKETRMPK
          **           **.   *    *    *    *

```

T QDTGNQSPLVSPSAMSSAAGKYHFSNLVSPTNLSQFNLPGPSMMRSNSIPAQDSSFDLY 1560  
T3 QDTAANSPFSSGSSVTSPSGTRFNFSQLASPTTVTQMSLSNPTMLRTHSLSNADGQYDPY  
T2 PESDDQSELPSPPALPMSLSAKGQLTNIVSPTAAT-----TPRITRSNSIPTHEAAFELY

T KLRRELVASQEKVATLTSQSLSANAHLVAAFEKGSLGNMTGRLQSLTMTAEGKESIELRE 1680  
 T3 KLRRELDASQEKVSLTTQLTANAHLVAAFEQSLGNMTIRIQLQSLTMTAEGKDSELNELRK  
 T2 KLRRELESSQEKVATLTSQSLSANANLVAAFEQSLVNMTRSRLRHLAETAEKDTELLDRE  
 \*\*\*\*\* . \*\*\*\*\* . \* . \* . \*\*\*\*\* . \* . \*\*\* . \* . \* . \* . \* . \* . \*

T TIEMPLKAQNSAAQAAIQGALNGPDHPPK-----DLRIRRQHSSESVSSINSATSHSS 1740  
 T3 TIELLKKQNAAAQAAINGVINTPELNCKGNGTAQOSADLRIRRQHSSDSVSSINSATSHSS  
 T2 TIDFLKKKNSEAQAVIQQGALNASETTPK-----ELRIKRQNSSDSISSLNSITSHSS  
 \*\*\*.\*\*\*.\*\*\*.\*\*\*.\*\*\*.\*\*\*.\*\*\*.\*\*\*.\*\*\*.\*\*\*.\*\*\*.\*\*\*.\*\*\*.\*\*\*\*\*

```

T  IGGSGNDADSKKKKKKNWL--RSSFKQAFGKKKSTKPPSSHS DIEELT--DSSLPA SPKL 1800
T3 VGSNIESDSKKKKRKNWVNELRSSFKQAFGKKKSPKSASSHSDIEEMT--DSSLPS PPKL
T2 IGGSSKDADAKKKKKKSWL--RSSFNKA FSIKKGP KSASSYSDIEEIATPDSSA PSSPKL
.** .-* .***** * *. **** * * * * * * * * * * * * * * * * * * * * * * * *

```

T	PHNAGDCGSASMKPSQSASAICECTEA-----	EAEIIILQLKSELRE	1860
T3	PHNGSTGSTPLLRNSHSNSLISECMDS-----	EAETVMQLRNELRD	
T2	QHGSTETASPSIKSSTS <del>SSV</del> GTDVTEGPAHPAPHTRLFHANE <del>EEE</del> EPEKKEVSELSELWE		

T SSSTSSSSSRQSLGLSLNNLNITEAVSSDILLDDAGDATGHKDG-RSVKIIIVSISKGYGR 1980  
T3 SISAS---PRQSMGLSQHSLNLTESTSLDMMLDDTGECSARKEGGRHVKIVVSFQEEMKW  
T2 PGSSALSSPRRSGLALTHSFGPSLADTDLSPMDGISTCGPKEE-VTLRVVVVRMPPQHII  
\* \* \* \* \*

09/914549

T AKDQKSQAYLIGSIGVSGKTKWDVLGVIRRLFKEYVFRIDTSTSGLSSDCIASYCID 2040  
T3 KEDSRPHLFLIGCIGVSGKTKWDVLGVRRLFKEYIIHVDPVSQQLGLNSDSVLGYSIGE  
T2 KGDLKQQEFFLGCSKVSGKVDWKMLDEAVFQVKDYISKMDPASTLGLSTESIHGYSISH  
\* . . . \* . \*\*\* \* . \*\* . . . \* . . \* . \*\*\* . . . \* . . \* .  
  
T LIRSHNLEVPELLPCGYLVGDNNIITVNLKGVEENSLSFVFDTLIPKPITQRYFNLLME 2100  
T3 IKRSNTSETPELLPCGYLVGENTTISVTVKGLAENSLSLVSFESLIPKPILQRYVSLLIE  
T2 VKRVLDAEPPPEMPPCRRGVNN--ISVSLKGLKEKCVDSLVFETLIPKPMMQHYISLLK  
. \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* .  
  
T HHRIILSGPSGTGKTYLANKLAEYVITKSGRKKTEDAIATFNVDHKSSKELQQYLANLAE 2160  
T3 HRRIILSGPSGTGKTYLANRLSEYIVLREGRELTDGVIATFNVDHKSSKELRQYLSNLAD  
T2 HRRVLVLSGPSGTGKTYLTNRLAELYVERSGREVTEGIVSTFNMHQSQCKDLQLYLSNLAN  
. \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* .  
  
T QCSADNNGVELPVVIILDNLHHVGSLSDIFNGFLNCKYNKCPYIIGTMNQGVSSSPNLEL 2220  
T3 QCNEENNNAVDMPPLVIILDNLHHVSSLGEIFNGLLNCKYHKCPYIIGTMNQATSSTPNQL  
T2 QIDRETGIGDVPLVILLDDLSEAGSISELVNGALTCKYHKCPYIIGTTNQPVKMTPNHGF  
. \* . . . \* . \* . \* . \* . . \* . . \* . \* . \* . \* . \* . \* . \* . \* .  
  
T HHNFRWVLCANHTEPVKGFLGRYLRRKLIEIEIERNIRNNDLVKIIDWIPKTWHHLNSFL 2280  
T3 HHNFRWVLCANHTEPVKGFLGRFLRRKLMETEISGRVRNMELVKIIDWIPKVWHHLNRFL  
T2 HLSFRMLTFSNNVEPANGFLVRYLRRKLVESDSDINANKEELLRVLDWVPKLYWHLHTFL  
. \* . \* . . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* .  
  
T ETHSSSDVTIGPRLFLPCPMVEGSRVWFMDLNYSLVPYILEAVREGLQMYGKRTPWED 2340  
T3 EAHSSSDVTIGPRLFLSCPIDVDGSRVWFMDLNYSIIPYLLEAVREGLQLYGRRAPWED  
T2 EKHSTSDFLIGPCFFLSCPIGIEDFRTWFIDLWNNSIIIPYLQEGAKDGIKVHGQKAAWED  
. \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* .  
  
T PSKWVLDTYPWSSATLPQESPALLQLRPEDVGYESCTSTKEATTSKHIPQTDTEGDPLMN 2400  
T3 PAKWVMDTYPWAASPOQHEWPPLLQLRPEDVGFDDGYSMMPREGSTSKQMPPSDAEGDPLMN  
T2 PVEWVRDTLPWPSAQQDQSK--LYHLPPPTVGPHSIASPPEDRTVKDSTPSSLSDDPLMA  
. \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* .  
  
T MLMKLQEAANYSSQTQSCDSES--TSHHEDILDSSLESTL  
T3 MLMRLQEAANYSSPQSYDSDSNSNSHHDDILDSSLESTL  
T2 MLLKLQEAANYIE--SPDRET-----ILDPNLQATL  
. \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* . \* .

Fig. 21 (cont'd 3)

09/914549

108 / 124

### Alignment of the T protein with the POM121 protein

10                  20                  30                  40                  50                  .60

T-Protein MDLSSEMNHRGKNPVSHKLEDQKKIYTDWANHYLAKSGHKRLIKDLQQDIADGVLLAEII  
POM121 -----

T-Protein QI IANEKVEDINGCPRSQSOMIENVDVCLSFLAARGVNVQGLSAEEIRNGNLKAILGLFFF  
 POM121 ---MSPAAAAADGGERRRP-----PLGVREGRGR-TRGCGGPAGAAALGLALLGLAL

T-Protein POM121	SLSRYKQQQHQQQYYQLVELQQRVTHASPPSEASQAKTQQDMQSRLPGPSRVPAGSS YLV-----P---AAAALAWLAVGASAAWGLSREPRGP * * * * *
---------------------	---

T-Protein SKVQGASNLLNRRSQFNSIDKNKPPNYANGNEKGEDPETRRMRTVKNIADLRQNLEETMS  
 M121 ---RGLSSFVRESR-----RHPRPALTASPLPAKSP-----VNGSLCBPRS

<b>T-Protein</b> <b>POM121</b>	<b>SLRGQTQISHSTLETTFDSTVTTEVNNGRTIPNLTSLRPTPMTWRLGQACPRLQAGDAPSGLA</b> <b>PLGGPDPAELLMGSGYLG-----KPGPPEPALPQD-PRDRPGRRPPSRS</b>
-----------------------------------	--

T-Protein GYPRSGTSRFIHTDPSRFMYTTPLRRRAAVSRLGNMSQIDMSEKASSDLDMSSEVDVGGYMP  
 POM121 PPSSSTAQRVHHVYP---ALPTPLLRSRR-----PPHRDCGPLS  
 \* \* \* \* \*\*\* \* . \* \* \*

T-Protein POM121	EELKKPEEDFD SHGDAGGKWKTVSSGLP EDPEKAGQKA SLVSQTGSWRRGMSAQGGAPS LHLDGQENKRRRHDSSG-----SGHSAFPLVANGVPAAFVPKPGSLKRSI LASQSSDDH
---------------------	--

T-Protein POM121 RQK-AGTSALKTPGKTDDAKASEKGK-APLKGSLLQRSPSDAGKSSGDEGKPPSGIGRS  
LNKRSRTSSVSSLTSTCTGGIPSSSRNAITSSYSSTRGVSQLWKRSG-PTSSPFSSPASS

T-Protein            TATSSFGFKKPSGVGSSAMITSSGATITSGSATLGKIPKSAAIGGKSNAGRKTSLDGSQNR  
 POM121            RSQTPERPAKTRSEEPCHQSSSAPLVTDKESPGEKVTDPATGKQQSLWTSPPTPGSSG

09/914549

109/124

T-Protein  
POM121

EPTKIGSGRSPVTVNQTDKEKEKVAVSDSESVLSGSPKSSPTSASACGAQGLRQPGSK  
WFNKVLEDKTDDASTPATDTSP---ATSPPFITLTL---P---TGVPAASPAPLAPSS-

T-Protein  
POM121

YPDIASTFRRLFGAKAGGKSASAPNTEGVKSSSVMPSPSTTLARQGSLESPSSGTGSMG  
----NPLLESLKMKMQESPAPSSSEPPE--AATVAAPSPPKTPSLLAPLVSP-----

T-Protein  
POM121

SAGGLSGSSSPLFNKPSDLTTDVISLSHSLASSPASVHSFTSGGLVWAANMSSSAGSKD  
----LTG---PLASTSSDSKPTTFGLASASSATPLTDTKAPGVSQALCVSTPAATAP

T-Protein  
POM121

TPSYQSMTSLHTSSESIDLPLSHHGSLSGLTTCHEVQSLLMRTGSVRSTLSESMQLDRN  
SP-----TPASTLFGMLSPPASSSLATPGPACASPMFKPIFPATPKSE---SDN

T-Protein  
POM121

TLPKKGLRYTPSSRQANQEEGKEWLRSHTGGLQDTGNQSPLVSPSAMSSAAGKYHFSN  
PLP-----TSSSAATTTPASTALPTTATAHTFKPIFESVEPFAAMP-----

T-Protein  
POM121

LVSPTNLSQLP GPSMMRSNSIPAQDSSFDLYDDSOLCGSATSLERPRAISHGSFRD  
LSPPFLSLKQTTPATTAATSAPLLTG-----L-----GTATST-----VATGTTAS

T-Protein  
POM121

SMEEVHGSSLVLVSSSTSSLYSTAEEKAHSEQIHKLRRELVASQEKVATLTSQLSANAHLV  
ASKPVFGFGVTTAASTASTIAS-----TSQSILFGGAPPVTASSAPALASIFQFGKPLA

T-Protein  
POM121

AAFEKSLGNMTGRLOSLTM TAEQKESELIELRETIEMLKAQNSAAQAAIQGALNGPDHPP  
PAASVAGTSFSQSLASSAQTAASNSS--GGFSCFGGTLLTSTSAPATTQOPTLTFNSNTVT

T-Protein  
POM121

KDLRIRRQHSSE-SVSSINSATSHSSIGSGNDADSKKKKKNWLRSSFKQAFGKKKSTK-  
PTFNIPFSASAKPALPTYPGANQPTFG-ATDGATKP---ALAPSFGSSFTFGNSVAS

T-Protein  
POM121

PPSSHSDIEELTDSSL PASPKLPHNAGDCGSASMKPQSASAICECTEAEAEIILQLKSE  
APSAAPAPAFAFGGAAQPAFGLKASASTFG---TPASTQPAFGSTTS-----VFSFGSA

T-Protein  
POM121

LREKELKLTDIRLEALSSAHLDQIREAMNRQNEIEILKAENDRLKAETGNTAKPTRPP  
TTS-----GFGAAAATTQTHSGS-----SSSLFGSSTPS-PF

T-Protein  
POM121

SESSSSTSSSSRQSLGLSLNNLNITEAVSSDILLDDAGDATGHKDGRSVKIIIVSISKY  
TFFGSAAPAGGG---GFGLSATPGTGSTSGTFSFGSGQSGT---TGTTSFGGSLSQNT

T-Protein  
POM121

GRAKDQKSQAYLIGSIGVSGKTKWDVLDGVIRRLFKEYVFRIDTSTS LGLSSDCIASYCI  
LGAPSQSS--PFAFSVGSTPESKP-----VFGGTSTPTFGQSAPAPG--V

09/914549

T-Protein  
POM121

GDLIRSHNLEVPPELLPCGYLVGDNNIITVNLKGVEENSLDSFVFDTLIPKPITQRYFNLL  
GTTGSSLSFGAPSTPAQGFVG-----VGPFPGAPSFSIGAGSKTPGARQRLQAR

\* \* \* \* . \* \* \* \* \* ..

T-Protein  
POM121

MEHHRIILSGPSGTGKTYLANKLAEYVITKSGRKKTEDAIATFNVVDHKSSKELQQYLANL  
RQHTRKK-----

\* \*

T-Protein  
POM121

AEQCSADNNGVELPVVIILDNLHHVGSLSDIFNGFLNCKYNKCPYIIGTMNQGVSSSPNL  
-----

T-Protein  
POM121

ELHENFRWVLCANHTEPVKGFLGRYLRRKLIEIEIERNIRNNDLVKIIDWIPKTWHHLNS  
-----

T-Protein  
POM121

FLETHSSSDVTIGPRLFLPCPMDEVGSRVWFMDLWNYSLVPLYILEAVREGLQMYGKRTPW  
-----

T-Protein  
POM121

EDPSKWVLDTYPWSSATLPQESPALLQLRPEDVGYESCTSTKEATTSKHIPQTDTEGDPL  
-----

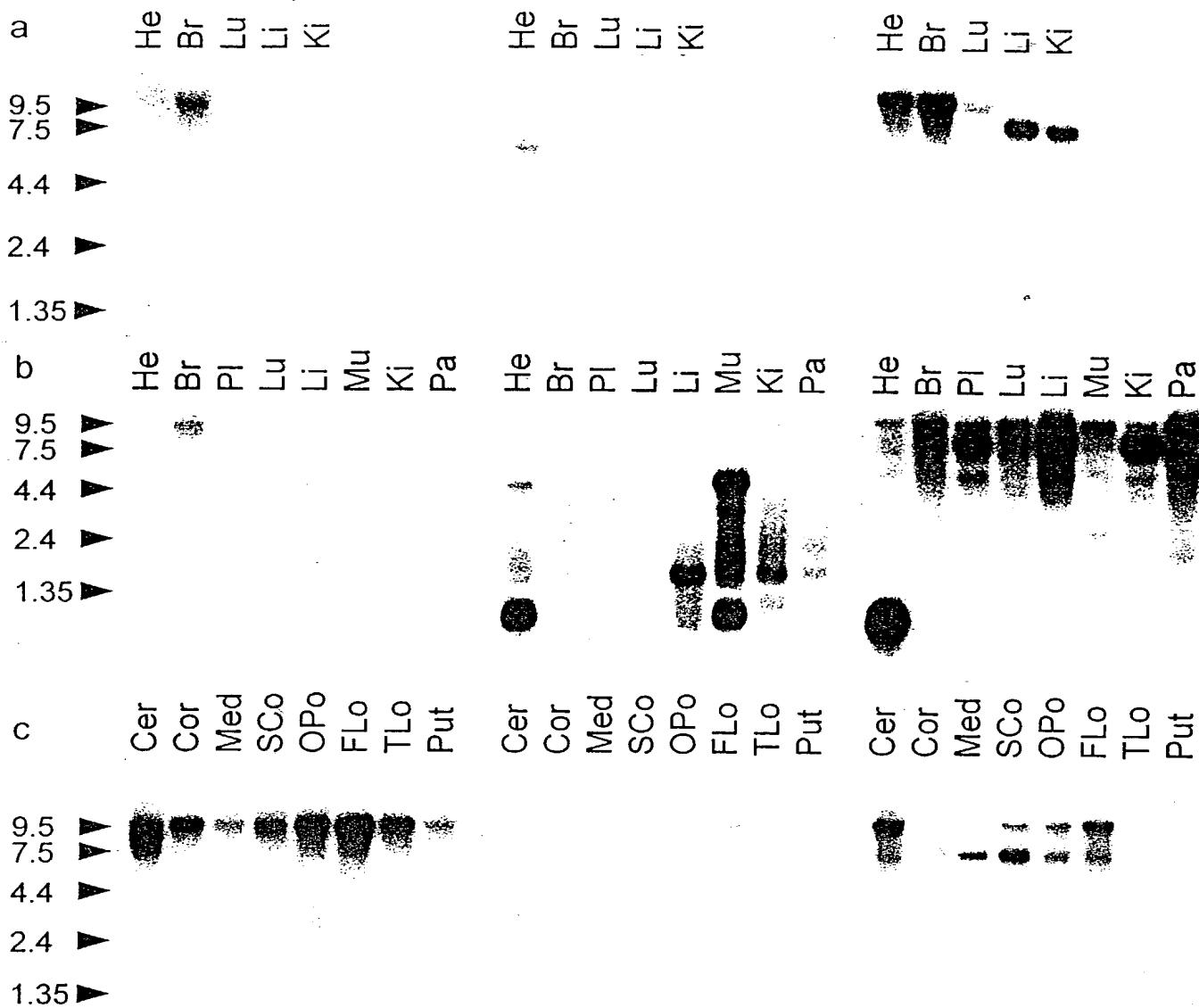
T-Protein  
POM121

MNMLMKLQEAANYSSTQSCDSESTSHHEDILDSSLESTL  
-----

Fig. 22 (cont'd 2)

09/914549

111/124



Expression of the T gene family.

**a** fetal tissue: left: T gene; middle: T2 gene; right: T3 gene.

He = heart; Br = brain; Lu = lungs; Li = liver; Ki = kidney

**b** adult tissue: left: T gene; middle: T2 gene; right: T3 gene.

He = heart; Br = brain; Pl = placenta; Lu = lungs; Li = liver; Mu = skeletal muscle; Ki = kidney; Pa = pancreas

**c** adult brain regions: left: T gene; middle: T2 gene; right: T3 gene.

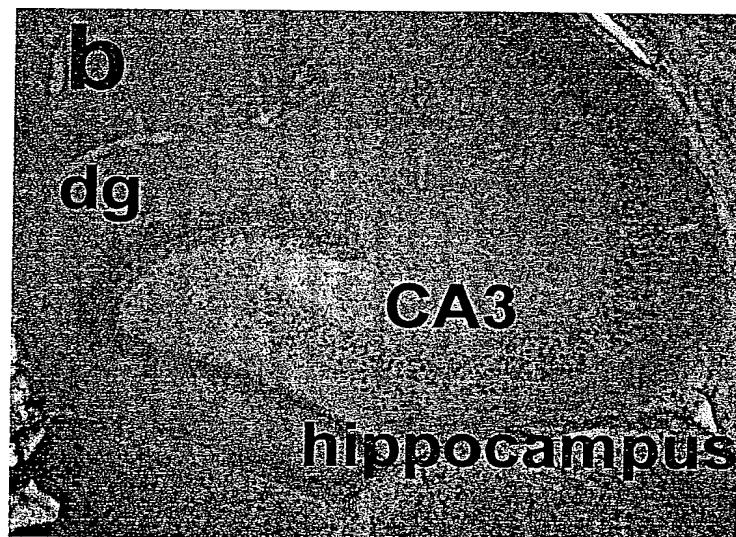
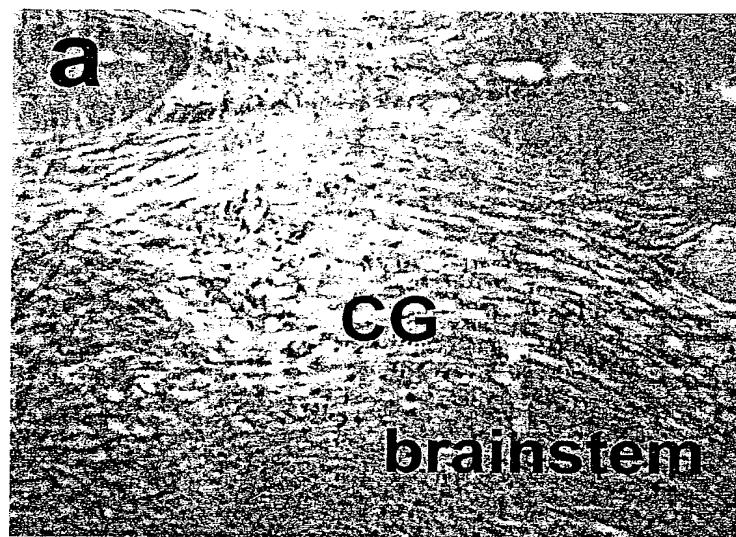
Cer = cerebellum; Cor = cerebral cortex; Med = medulla; Sco = spinal cord; Opo = occipital pole; Flo = frontal lobe; Tlo = temporal lobe; Put = putamen

Fig. 23

09/914549

112/124

Fig. 24



09/914549

113/124

Fig. 24

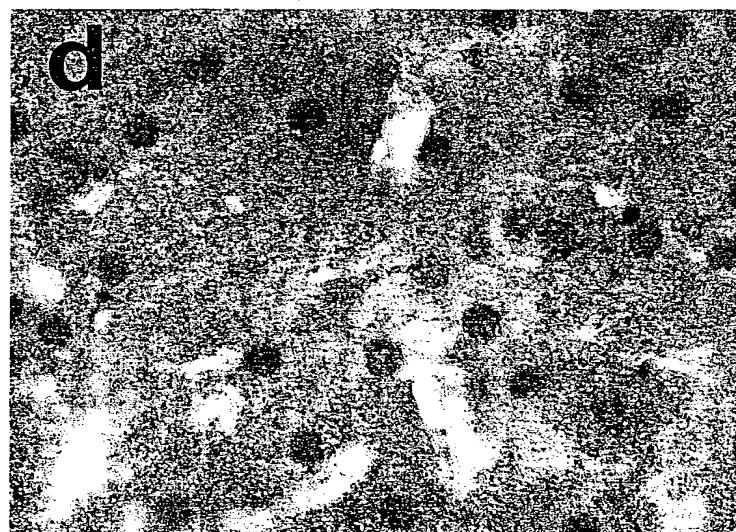
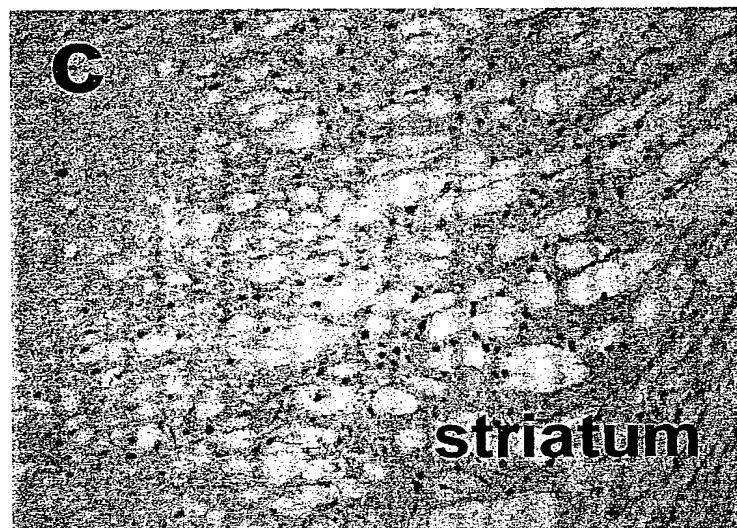


Fig. 24

114/124

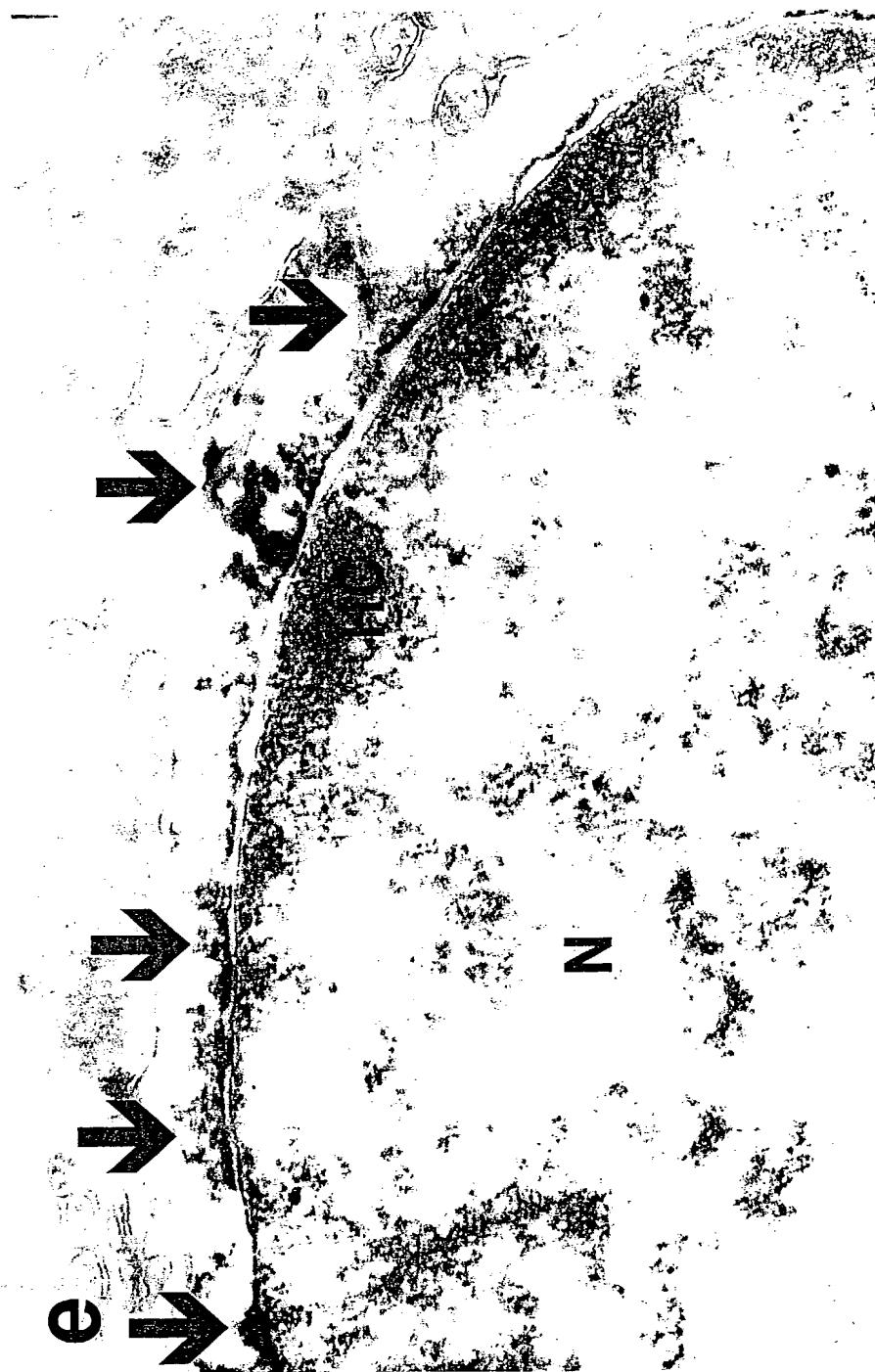


Figure legend of immunohisto and electron microscopy:

a = brain stem. CG central grey = central grey of the brain stem

b = hippocampus. dg = dental gyrus; CA3 cornu ammonis 3, both subregions of the hippocampus formation

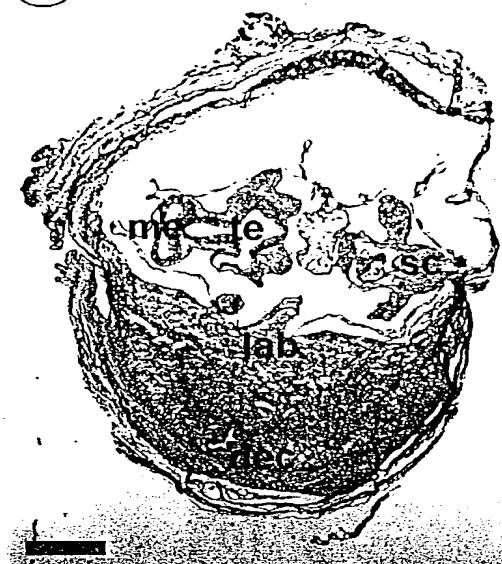
c = electronmicroscopic picture. N = nucleus, Hc heterochromatin

09/914549

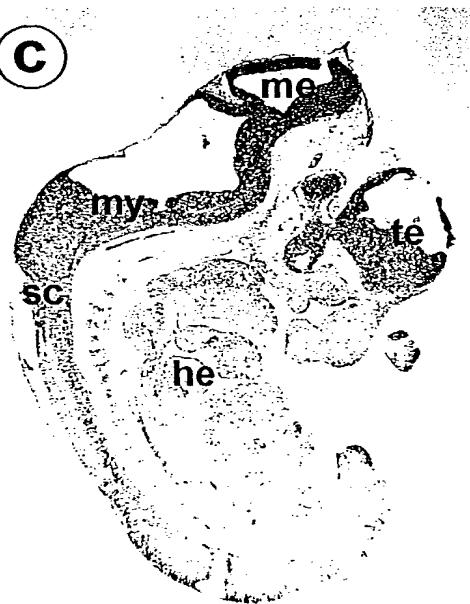
Fig. 25

115/124

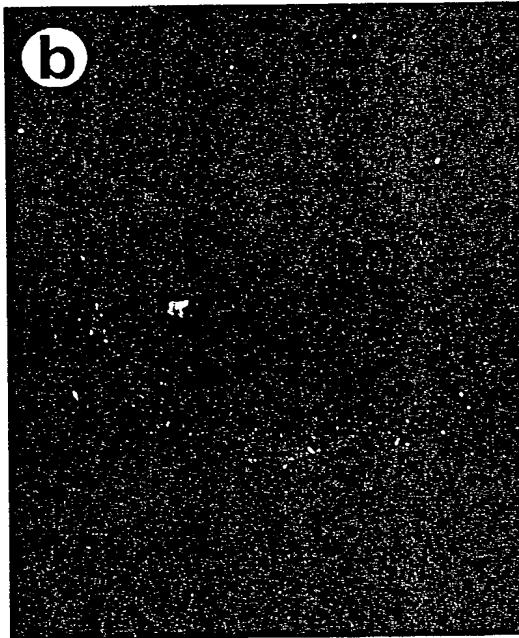
(a)



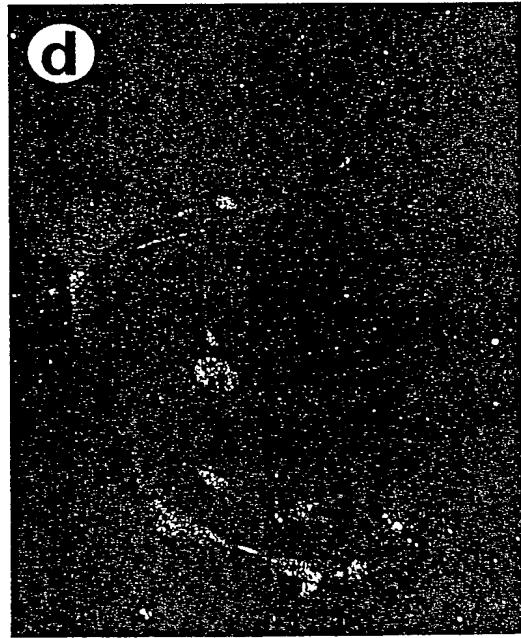
(c)



(b)



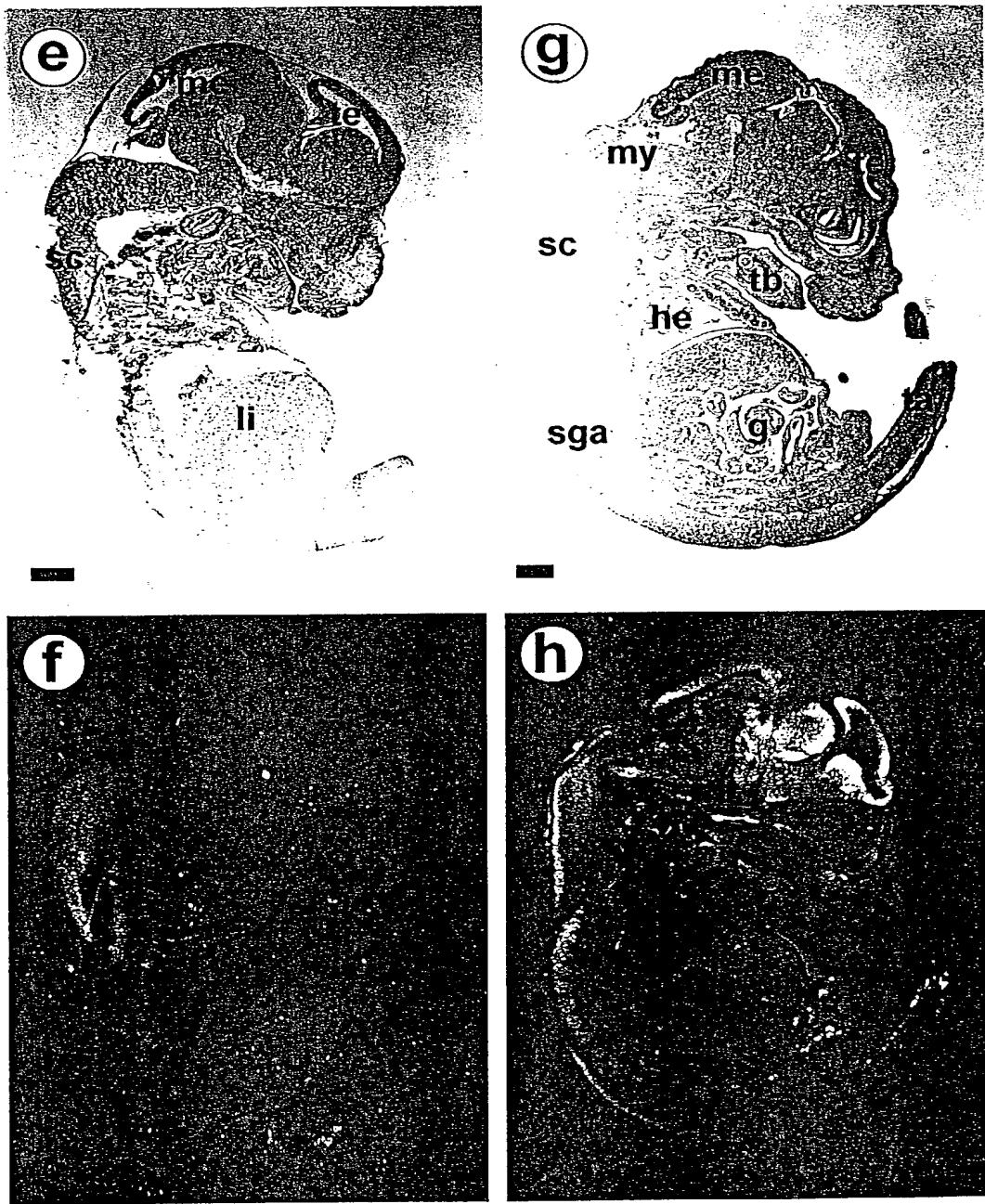
(d)



09/914549

116/124

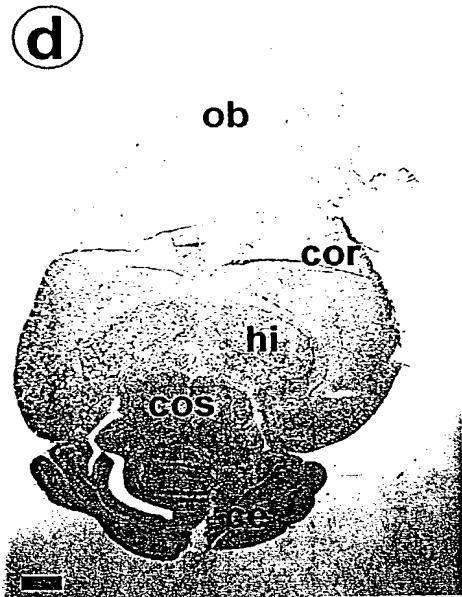
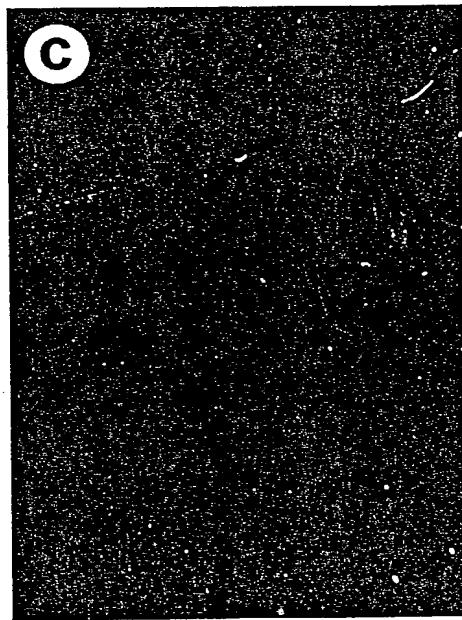
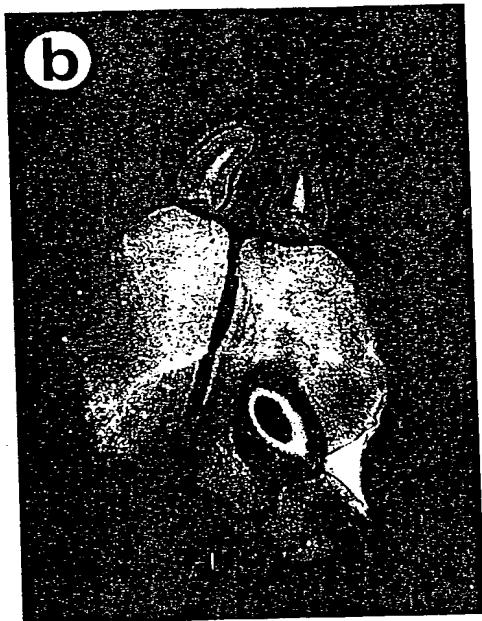
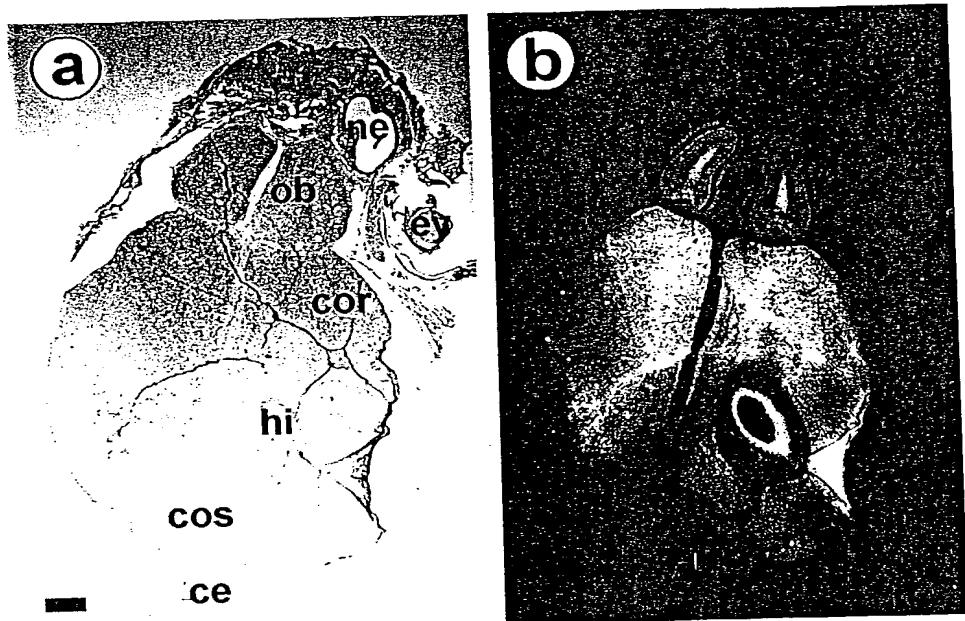
Fig. 25



09/914549

Fig. 26

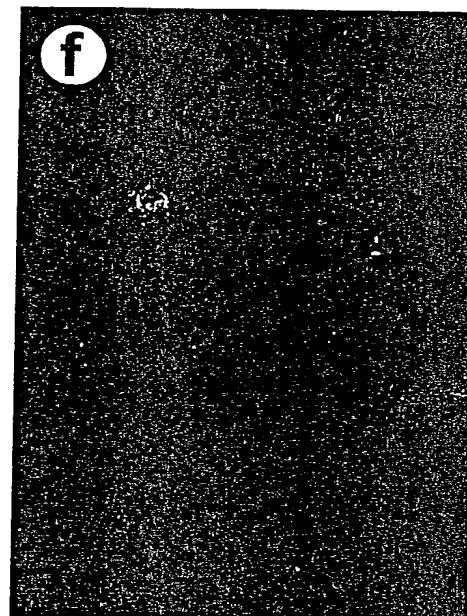
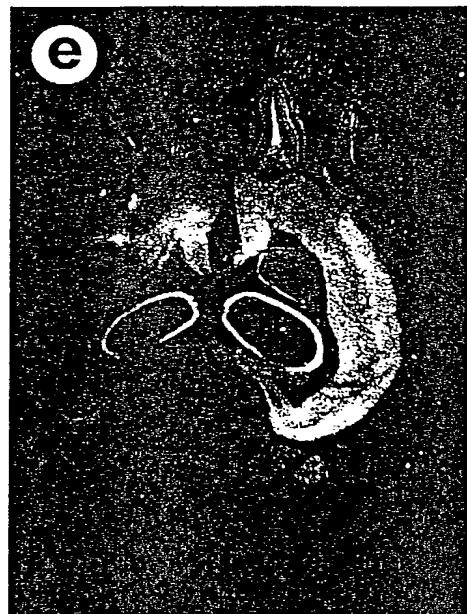
117/124



09/914549

118/124

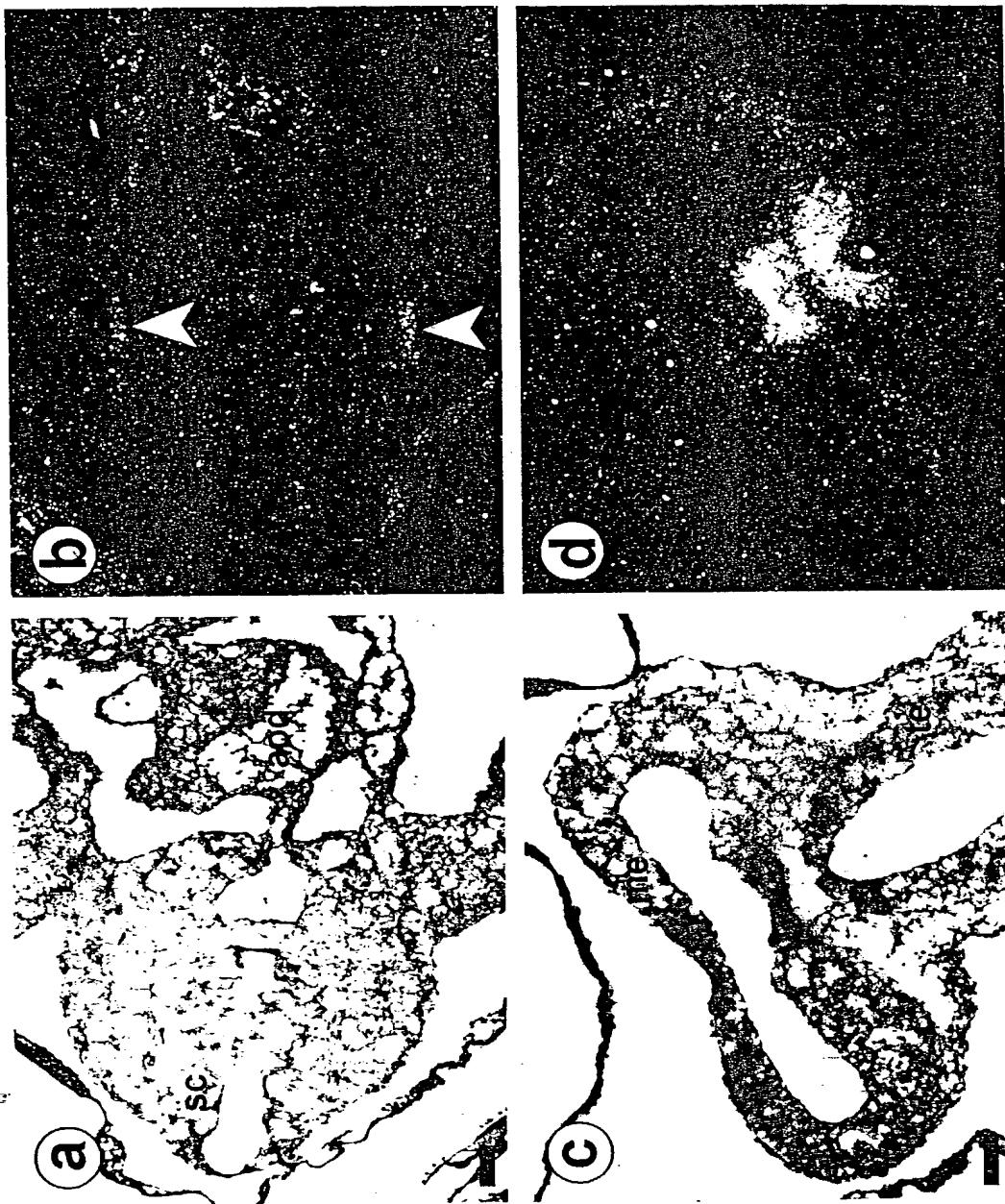
Fig. 26



097914549

Fig. 27

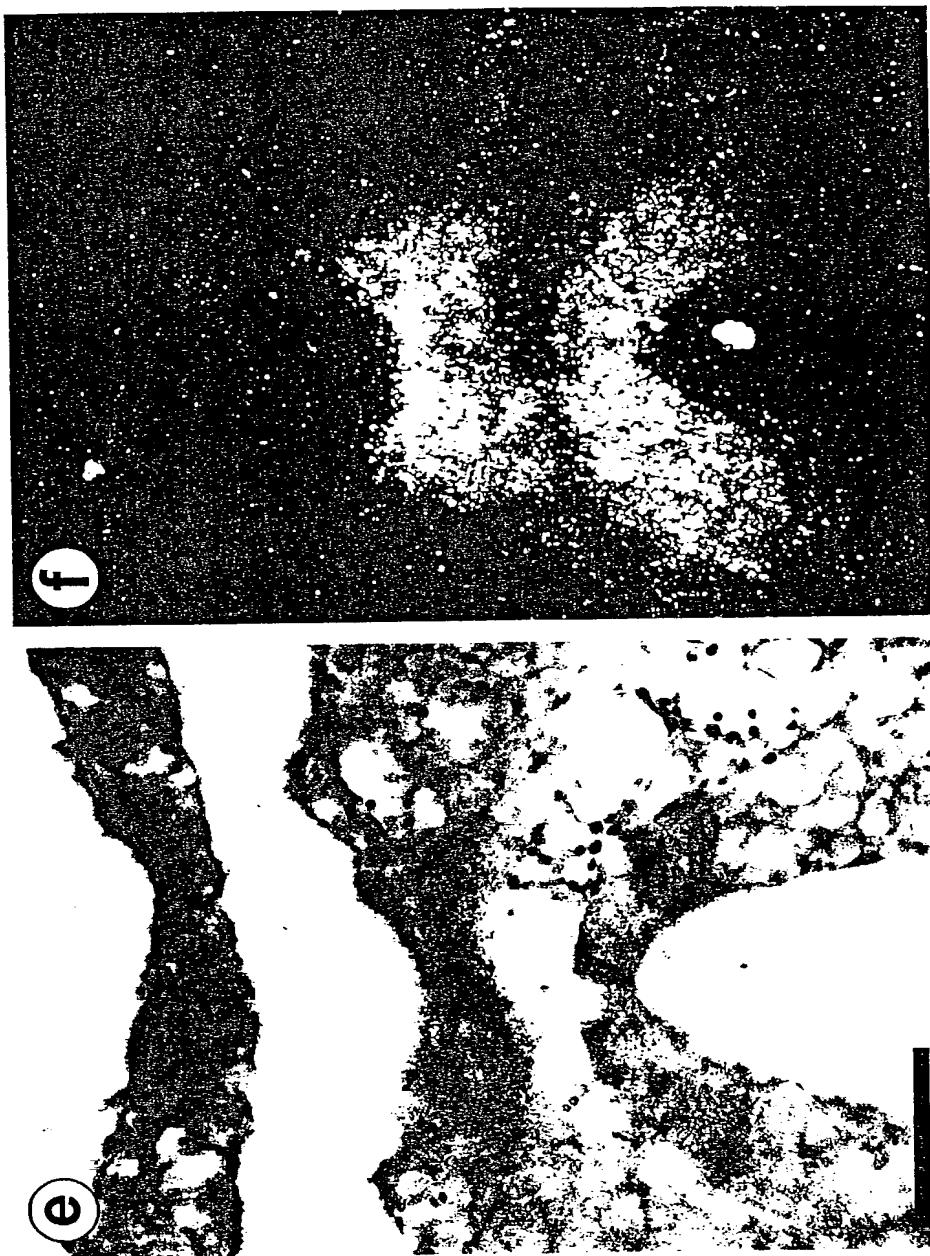
119/124



09/914549

120/124

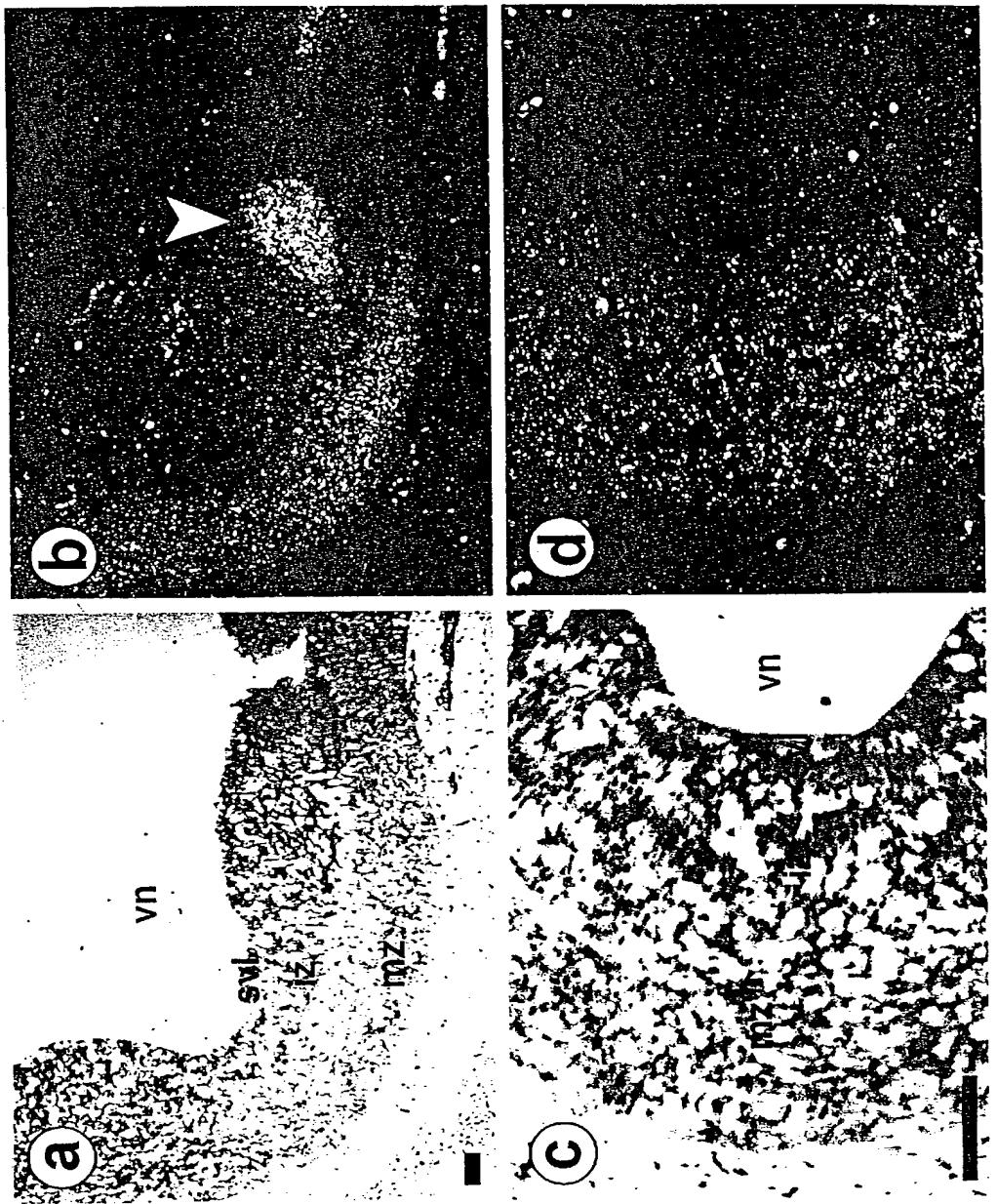
Fig. 27



09/914549

121/124

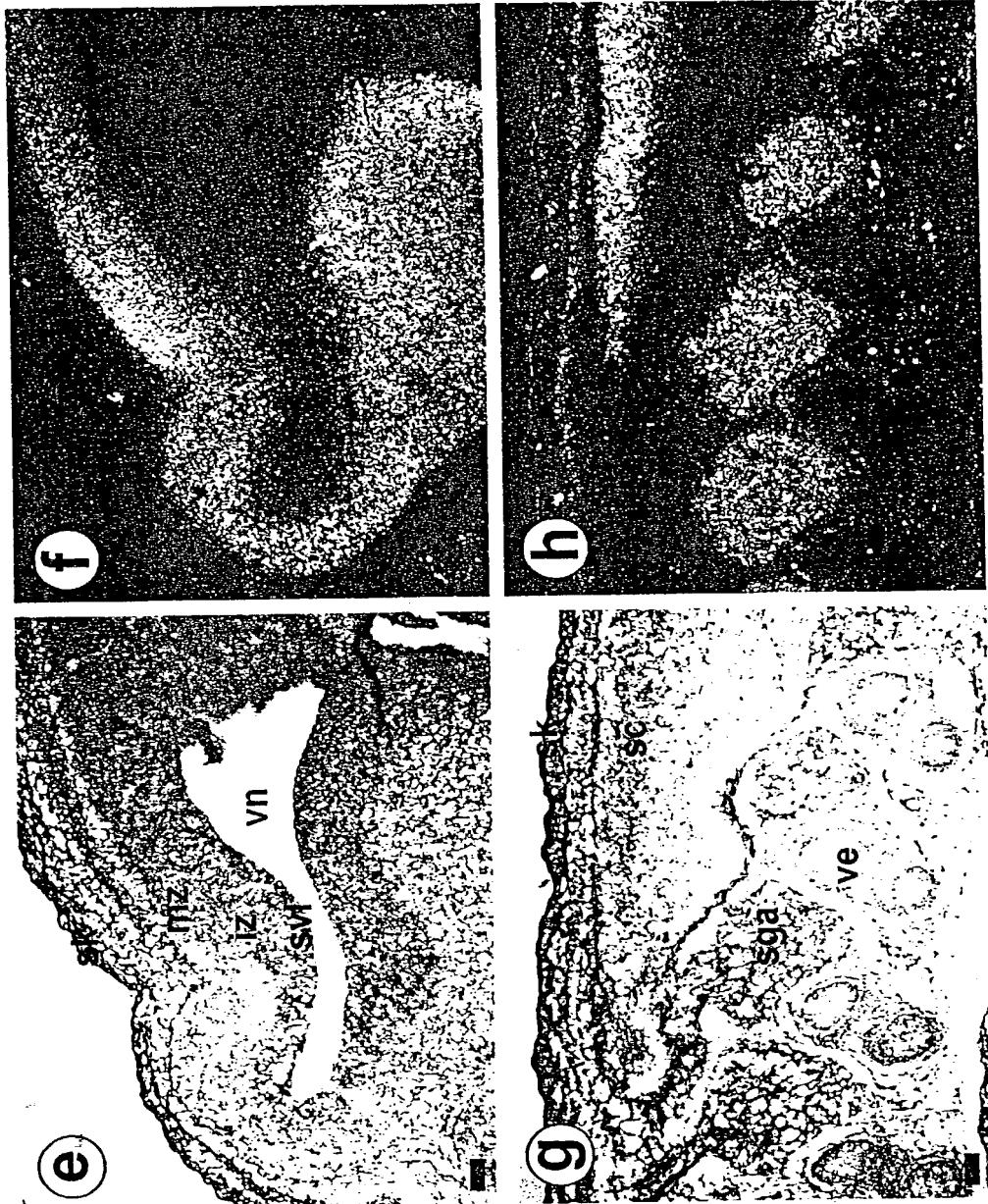
Fig. 28



09/914549

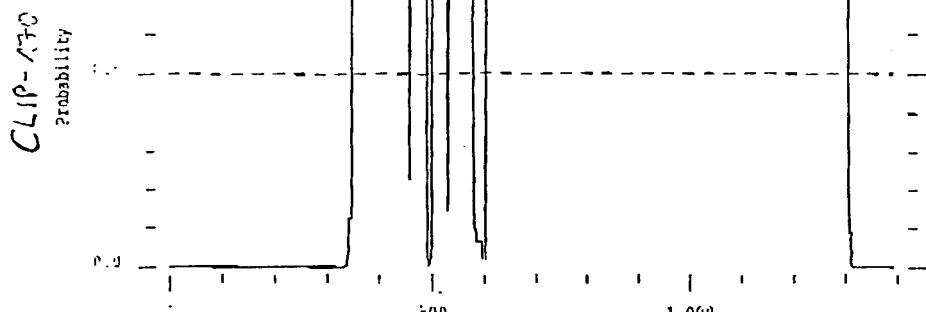
Fig. 28

122/124

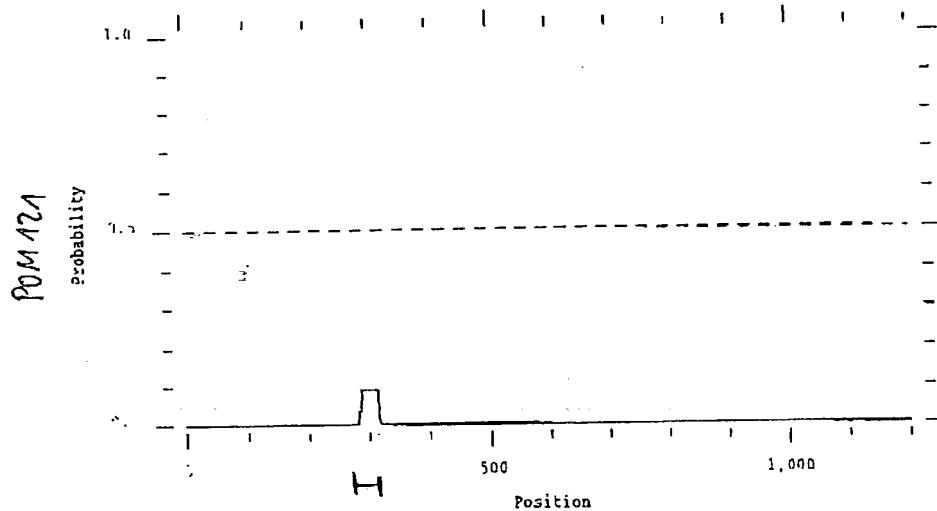
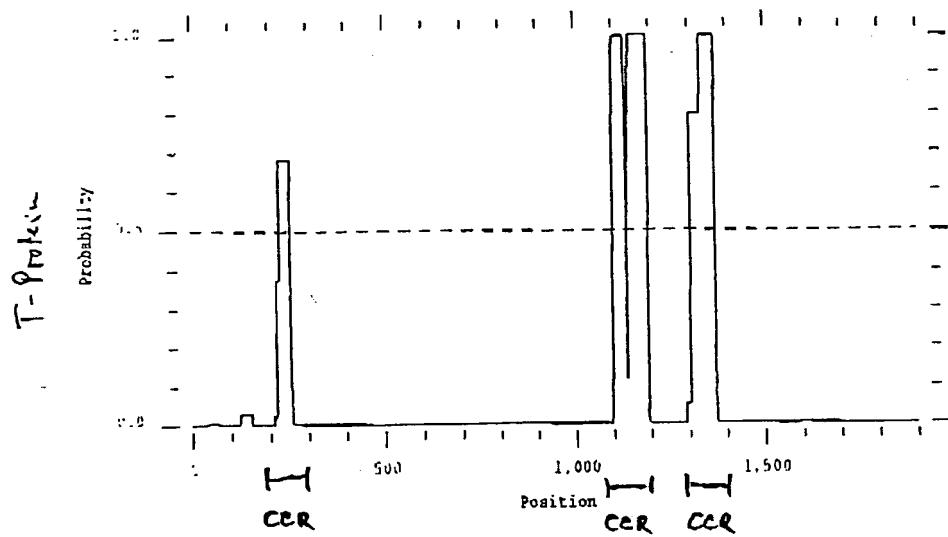


09/914549  
123/124

Fig. 29



Coiled-Coil  
Region (CCR)

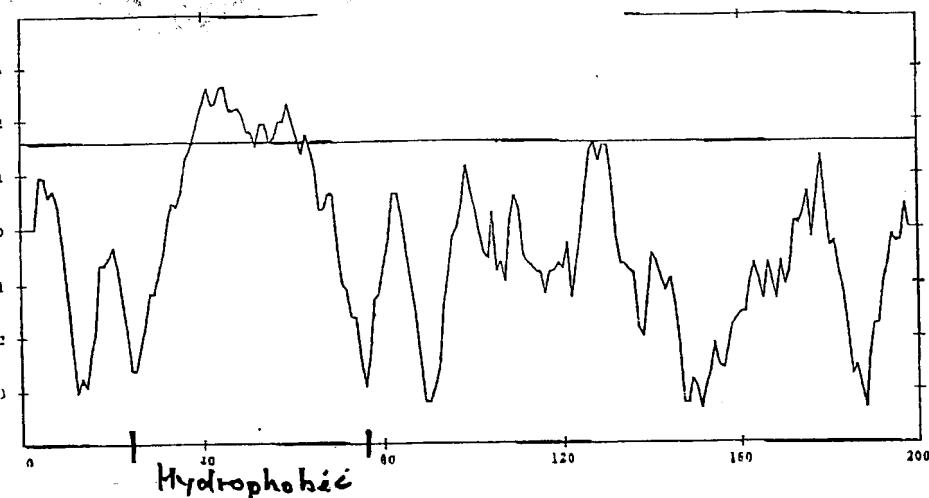


09/914549

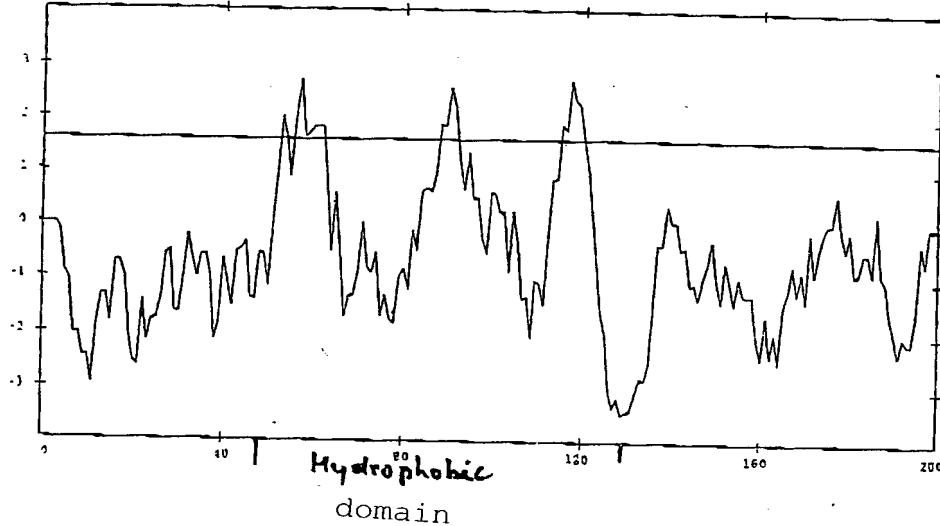
124/124

Fig. 30

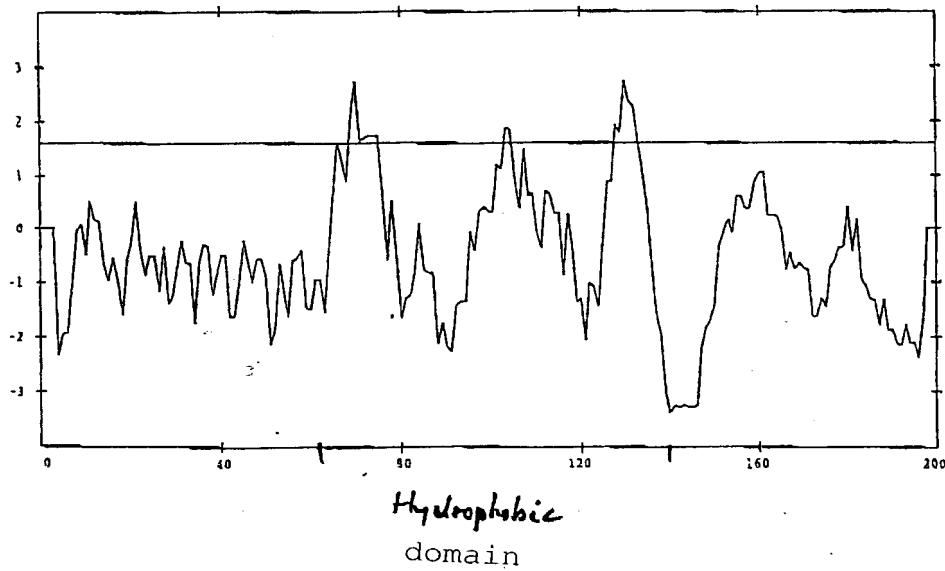
POM124



T-Protein



T3-Protein



## PATENT APPLICATION

DECLARATION AND POWER OF ATTORNEY  
FOR PATENT APPLICATION

ATTORNEY DOCKET NO. 4121-129

As a below named inventor, I hereby declare that:

My residence/post office address and citizenship are as stated below next to my name;

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter, which is claimed and for which a patent is sought on the invention entitled:

PROTEIN (TP) THAT IS INVOLVED IN THE DEVELOPMENT OF THE NERVOUS SYSTEM

the specification of which is attached hereto unless the following box is checked:

was filed on August 24, 2001 as US Application Serial No. 09/914,549 or PCT International Application Number \_\_\_\_\_ and was amended on \_\_\_\_\_ (if applicable).

I hereby state that I have reviewed and understood the contents of the above-identified specification, including the claims, as amended by any amendment(s) referred to above. I acknowledge the duty to disclose all information which is material to patentability as defined in 37 CFR 1.56.

## Foreign Application(s) and/or Claim of Foreign Priority

I hereby claim foreign priority benefits under Title 35, United States Code Section 119(a-d) or 365(b) of any foreign application(s) for patent or inventor(s) certificate, or 365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below any foreign application for patent or inventor(s) certificate having a filing date before that of the application on which priority is claimed:

COUNTRY	APPLICATION NUMBER	DATE FILED	PRIORITY CLAIMED UNDER 35 U.S.C. 119
Germany	199 08 423.8	26 February 1999	YES: <input checked="" type="checkbox"/> NO: _____
PCT	PCT/DE00/00583	28 February 2000	YES: <input checked="" type="checkbox"/> NO: _____

## Provisional Application

I hereby claim the benefit under Title 35, United States Code Section 119(e) of any United States provisional application(s) listed below:

## U.S. Priority Claim

I hereby claim the benefit under Title 35, United States Code, Section 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claim of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code Section 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, Section 1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

APPLICATION SERIAL NUMBER	FILING DATE	STATUS(patented/pending/abandoned)

## POWER OF ATTORNEY:

As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) listed below to prosecute this application and transact all business in the Patent and Trademark Office connected therewith.

Steven J. Hultquist, Reg. No. 28,021

Marianne Fulerer, Reg. No. 39,983

## Send Correspondence to:

## Direct Telephone Calls To:

Steven J. Hultquist  
Intellectual Property/Technology Law  
P.O. Box 14329  
Research Triangle Park, NC 27709

Steven J. Hultquist  
(919) 419-9350

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full Name of Inventor: Annemarie PoustkaCitizenship: AustrianResidence: Wernerstrasse 36, D-69120 Heidelberg, GermanyExadenburger Str. 41AT 20.12.01Post Office Address: Same30. Okt. 01Inventor's Signature: Annemarie Poustka

Date

**DECLARATION AND POWER OF ATTORNEY  
FOR PATENT APPLICATION**

**ATTORNEY DOCKET NO. 4121-129**

*200*  
Full Name of Inventor: Johannes Coy

Citizenship: German

Residence: In den Schwarzen Garten 1, D-63762 Grossostheim, Germany

Post Office Address: Same

Johannes Coy  
Inventor's Signature

Date

*30.10.01*

09/09/2001  
518 Rec'd PCT/PTO 24 AUG 2001

SEQUENCE LISTING

<110> Poustka, et al.

<120> Protein (TP) That is Involved in the Development of the Nervous System

<130> 4121-129

<140>

<141>

<150> PCT/DE00/00583  
DE 199 08 423.8

<151> 2000-02-28  
1999-02-26

<160> 39

<170> PatentIn Ver. 2.0

<210> 1  
<211> 242  
<212> DNA  
<213> Homo sapiens

<220>  
<221> CDS  
<223> (3)..(242)

<400> 1

ag cga gtt act cac gct tcc cct cca tcg gaa gcc agc cag gcc aaa 47  
Arg Val Thr His Ala Ser Pro Pro Ser Glu Ala Ser Gln Ala Lys  
1 5 10 15

acc cag caa gat atg cag tcc agt ctg gca gcc aga tat gca act cag 95  
Thr Gln Gln Asp Met Gln Ser Ser Leu Ala Ala Arg Tyr Ala Thr Gln  
20 25 30

tct aat cac agt gga att gca acc agt caa aaa aag cct act agg ctt 143  
Ser Asn His Ser Gly Ile Ala Thr Ser Gln Lys Lys Pro Thr Arg Leu  
35 40 45

cca ggg ccc tct agg gtg cct gct gca gga agc agc aag gtc cag 191  
Pro Gly Pro Ser Arg Val Pro Ala Ala Gly Ser Ser Ser Lys Val Gln  
50 55 60

gga gcc tct aat tta aat agg aga agt cag agc ttt aac agc att gac 239  
Gly Ala Ser Asn Leu Asn Arg Arg Ser Gln Ser Phe Asn Ser Ile Asp  
65 70 75

aaa 242  
Lys

80

<210> 2  
<211> 80  
<212> PRT  
<213> Homo sapiens

<400> 2

Arg Val Thr His Ala Ser Pro Pro Ser Glu Ala Ser Gln Ala Lys Thr  
1 5 10 15

Gln Gln Asp Met Gln Ser Ser Leu Ala Ala Arg Tyr Ala Thr Gln Ser  
20 25 30

Asn His Ser Gly Ile Ala Thr Ser Gln Lys Lys Pro Thr Arg Leu Pro  
35 40 45

Gly Pro Ser Arg Val Pro Ala Ala Gly Ser Ser Ser Lys Val Gln Gly  
50 55 60

Ala Ser Asn Leu Asn Arg Arg Ser Gln Ser Phe Asn Ser Ile Asp Lys  
65 70 75 80

<210> 3  
<211> 159  
<212> DNA  
<213> Homo sapiens

<220>  
<221> CDS  
<223> (1)..(159)

<400> 3

ggc act cac gag gtc cag agc ctg ctc atg aga acg ggt agt gtg aga 48  
Gly Thr His Glu Val Gln Ser Leu Leu Met Arg Thr Gly Ser Val Arg  
5 10 15

tct act ctc tca gaa aga tat acc cca tca tct cgg cag gcc aac caa 96  
Ser Thr Leu Ser Glu Arg Tyr Thr Pro Ser Ser Arg Gln Ala Asn Gln  
20 25 30

gaa gag ggc aaa gag tgg ttg cgt tct cat tct act gga qgg ctt cag 144  
Glu Glu Gly Lys Glu Trp Leu Arg Ser His Ser Thr Gly Gly Leu Gln  
35 40 45

gac act ggc aac cag 159  
Asp Thr Gly Asn Gln  
50

<210> 4  
<211> 53  
<212> PRT  
<213> Homo sapiens

<400> 4

Gly Thr His Glu Val Gln Ser Leu Leu Met Arg Thr Gly Ser Val Arg  
1 5 10 15

Ser Thr Leu Ser Glu Arg Tyr Thr Pro Ser Ser Arg Gln Ala Asn Gln  
20 25 30

Glu Glu Gly Lys Glu Trp Leu Arg Ser His Ser Thr Gly Gly Leu Gln  
35 40 45

Asp Thr Gly Asn Gln  
50

<210> 5  
<211> 2461  
<212> DNA  
<213> Homo sapiens

<220>  
<221> CDS  
<223> (2)..(1627)

<400> 5

g gat cag ctt cgg gag acc atg cac aac atg cag ttg gag gtg gac 46  
Asp Gln Leu Arg Glu Thr Met His Asn Met Gln Leu Glu Val Asp

ctg ctg aaa gca gag aat gac cga ctg aag gta gcc cca ggc ccc tca 94  
Leu Leu Lys Ala Glu Asn Asp Arg Leu Lys Val Ala Pro Gly Pro Ser

tca ggc tcc act cca ggg cag gtc cct gga tca tct gca tta tct tcc 142  
Ser Gly Ser Thr Pro Gly Gln Val Pro Gly Ser Ser Ala Leu Ser Ser

cca cgc cgc tcc cta ggc ctg gca ctc acc cat tcc ttc ggc ccc agt 190  
Pro Arg Arg Ser Leu Gly Leu Ala Leu Thr His Ser Phe Gly Pro Ser

ctt gca gac aca gac ctg tca ccc atg gat ggc atc agt act tgt ggt 238  
Leu Ala Asp Thr Asp Leu Ser Pro Met Asp Gly Ile Ser Thr Cys Gly

cca aag gag gaa gtg acc ctc cgg gtg gtg agg atg ccc ccg cag 286  
Pro Lys Glu Val Thr Leu Arg Val Val Arg Met Pro Pro Gln

cac atc atc aaa ggg gac ttg aag cag cag gaa ttc ttc ctg ggc tgt	334
His Ile Ile Lys Gly Asp Leu Lys Gln Gln Glu Phe Phe Leu Gly Cys	
agc aag gtc agt gga aaa gtt gac tgg aag atg ctg gat gaa gct gtt	382
Ser Lys Val Ser Gly Lys Val Asp Trp Lys Met Leu Asp Glu Ala Val	
ttc caa gtg ttc aag gac tat att tct aaa atg gac cca gcc tct acc	430
Phe Gln Val Phe Lys Asp Tyr Ile Ser Lys Met Asp Pro Ala Ser Thr	
ctg gga cta agc act gag tcc atc cat ggc tac agc atc agc cac gtg	478
Leu Gly Leu Ser Thr Glu Ser Ile His Gly Tyr Ser Ile Ser His Val	
aaa cga gtg ttg gat gca gag ccc ccc gag atg cct cct tgc cgt cga	526
Lys Arg Val Leu Asp Ala Glu Pro Pro Glu Met Pro Pro Cys Arg Arg	
ggt gtc aat aac ata tca gtc tcc ctc aaa ggt ctg aag gag aaa tgc	574
Gly Val Asn Asn Ile Ser Val Ser Leu Lys Glu Leu Lys Glu Lys Cys	
gtc gac agc ctg gtg ttc gag acg ctg atc ccc aag ccg atg atg cag	622
Val Asp Ser Leu Val Phe Glu Thr Leu Ile Pro Lys Pro Met Met Gln	
cac tac ata agc ctc ctg ctg aag cac cgg cgc ctc gtc ctc tcg ggc	670
His Tyr Ile Ser Leu Leu Lys His Arg Arg Leu Val Leu Ser Gly	
ccc agc ggc acg ggc aag acc tac ctg acc aat cgc ttg gcc gag tac	718
Pro Ser Gly Thr Gly Lys Thr Tyr Leu Thr Asn Arg Leu Ala Glu Tyr	
ctg gtg gag cgc tct ggc cgt gag gtc aca gag ggc atc gtc agc acc	766
Leu Val Glu Arg Ser Gly Arg Glu Val Thr Glu Gly Ile Val Ser Thr	
ttc aac atg cac cag cag tct tgc aag gat ctgcaa ctg tat ctt tcc	814
Phe Asn Met His Gln Gln Ser Cys Lys Asp Leu Gln Leu Tyr Leu Ser	
aac cta gcc aac cag ata gac cgg gaa aca gga att ggg gat gtg ccc	862
Asn Leu Ala Asn Gln Ile Asp Arg Glu Thr Gly Ile Gly Asp Val Pro	
ctg gtg att cta ttg gat gac ctg agt gaa gca ggc tcc atc agt gag	910
Leu Val Ile Leu Leu Asp Asp Leu Ser Glu Ala Gly Ser Ile Ser Glu	
ttg gtc aat ggg gcc ctc acc tgc aag tat cat aaa tgt ccc tat att	958
Leu Val Asn Gly Ala Leu Thr Cys Lys Tyr His Lys Cys Pro Tyr Ile	

ata ggt acc acc aat cag cct gta aaa atg aca ccc aac cat ggc ttg Ile Gly Thr Thr Asn Gln Pro Val Lys Met Thr Pro Asn His Gly Leu	1006
cac ttg agc ttc agg atg ttg acc ttc tcc aac aac gtg gag cca gcc His Leu Ser Phe Arg Met Leu Thr Phe Ser Asn Asn Val Glu Pro Ala	1054
aat ggc ttc ctg gtt cgt tac ctg agg agg aag ctg gta gag tca gac Asn Gly Phe Leu Val Arg Tyr Leu Arg Arg Lys Leu Val Glu Ser Asp	1102
agc gac atc aat gcc aac aag gaa gag ctg ctt cg <sup>g</sup> gtg ctc gac tgg Ser Asp Ile Asn Ala Asn Lys Glu Glu Leu Leu Arg Val Leu Asp Trp	1150
gta ccc aag ctg tgg tat cat ctc cac acc ttc ctt gag aag cac agc Val Pro Lys Leu Trp Tyr His Leu His Thr Phe Leu Glu Lys His Ser	1198
acc tca gac ttc ctc atc ggc cct tgc ttc ttt ctg tcg tgt ccc att Thr Ser Asp Phe Leu Ile Gly Pro Cys Phe Leu Ser Cys Pro Ile	1246
ggc att gag gac ttc cg <sup>g</sup> acc tgg ttc att gac ctg tgg aac aac tct Gly Ile Glu Asp Phe Arg Thr Trp Phe Ile Asp Leu Trp Asn Asn Ser	1294
atc att ccc tat cta cag gaa gga gcc aag gat ggg ata aag gtc cat Ile Ile Pro Tyr Leu Gln Glu Gly Ala Lys Asp Gly Ile Lys Val His	1342
gga cag aaa gct tgg gag gac cca gtg gaa tgg gtc cg <sup>g</sup> gac aca Gly Gln Lys Ala Ala Trp Glu Asp Pro Val Glu Trp Val Arg Asp Thr	1390
ctt ccc tgg cca tca gcc caa caa gac caa tca aag ctg tac cac ctg Leu Pro Trp Pro Ser Ala Gln Gln Asp Gln Ser Lys Leu Tyr His Leu	1438
ccc cca ccc acc gtg ggc cct cac agc att gcc tca cct ccc gag gat Pro Pro Pro Thr Val Gly Pro His Ser Ile Ala Ser Pro Pro Glu Asp	1486
agg aca gtc aaa gac agc acc cca agt tct ctg gac tca gat cct ctg Arg Thr Val Lys Asp Ser Thr Pro Ser Ser Leu Asp Ser Asp Pro Leu	1534
atg gcc atg ctg ctg aaa ctt caa gaa gct gcc aac tac att gag tct Met Ala Met Leu Leu Lys Leu Gln Glu Ala Ala Asn Tyr Ile Glu Ser	1582
cca gat cga gaa acc atc ctg gac ccc aac ctt cag gca aca ctt Pro Asp Arg Glu Thr Ile Leu Asp Pro Asn Leu Gln Ala Thr Leu	1627

taagggttcg gcaatcactg tcaccccccgg acagcagaac gctggcatca gctatcttag	1687
ctcctcctct cccctctcct ctttcagagc actggctctc cagccccagg aggagaacag	1747
gagggaggag gagatgaaag aggaggaca gggttcttggt gctgtacctt tgagaacttc	1807
ctaggaagga atgggtgggt ggcgtttggg aacttgtgcc ccctaaacac atttactggc	1867
ctcctctaata gactttgggg aaaagatgat tctgggtctt tcccttgact tcttgggtca	1927
attacaaact cctgggcttt ctggggaggg gttcagaaaa catcaaaaca ctgcagcagt	1987
tcctaaatga ttctcacaag caaccctgag agagacagtc ttgtgaggg aatctgggg	2047
aggcaggaag ctcctcagat tttctcacag acccttccca attccatcac cactgccaac	2107
aactcctccc ccagagatct ggctggagcc cagaaaaaga agcatgttgt ttaaaaaatg	2167
tttaaatcaa tctgtaaaag gtaaaaatga aaaacaaaaa caagcaaaca aacaaaaaac	2227
aatggaaaag atgaagctgg agagagagga accagttgcc aaggttagaga gctgcccgt	2287
cctgcccctct ggatgacata ggggacatca acaagacggc tgccaacctg agaagtcacc	2347
aaaccacaaa aataaccta cagccttcag ggaaagacta ccagctctgt ctttctaccc	2407
tctaatttaa caatgcataa gagtcaataa accctacttt ttaaaaaaaaaaaa	2461

<210> 6  
<211> 542  
<212> PRT  
<213> Homo sapiens

<400> 6

Asp Gln Leu Arg Glu Thr Met His Asn Met Gln Leu Glu Val Asp Leu			
1	5	10	15
Leu Lys Ala Glu Asn Asp Arg Leu Lys Val Ala Pro Gly Pro Ser Ser			
20	25	30	
Gly Ser Thr Pro Gly Gln Val Pro Gly Ser Ser Ala Leu Ser Ser Pro			
35	40	45	
Arg Arg Ser Leu Gly Leu Ala Leu Thr His Ser Phe Gly Pro Ser Leu			
50	55	60	
Ala Asp Thr Asp Leu Ser Pro Met Asp Gly Ile Ser Thr Cys Gly Pro			
65	70	75	80
{ Lys Glu Glu Val Thr Leu Arg Val Val Val Arg Met Pro Pro Gln His			
85	90	95	
Ile Ile Lys Gly Asp Leu Lys Gln Gln Glu Phe Phe Leu Gly Cys Ser			

100	105	110
Lys Val Ser Gly Lys Val Asp Trp	Lys Met Leu Asp Glu Ala Val Phe	
115	120	125
Gln Val Phe Lys Asp Tyr Ile Ser Lys Met Asp Pro Ala Ser Thr Leu		
130	135	140
Gly Leu Ser Thr Glu Ser Ile His Gly Tyr Ser Ile Ser His Val Lys		
145	150	155
Arg Val Leu Asp Ala Glu Pro Pro Glu Met Pro Pro Cys Arg Arg Gly		
165	170	175
Val Asn Asn Ile Ser Val Ser Leu Lys Gly Leu Lys Glu Lys Cys Val		
180	185	190
Asp Ser Leu Val Phe Glu Thr Leu Ile Pro Lys Pro Met Met Gln His		
195	200	205
Tyr Ile Ser Leu Leu Leu Lys His Arg Arg Leu Val Leu Ser Gly Pro		
210	215	220
Ser Gly Thr Gly Lys Thr Tyr Leu Thr Asn Arg Leu Ala Glu Tyr Leu		
225	230	235
240		
Val Glu Arg Ser Gly Arg Glu Val Thr Glu Gly Ile Val Ser Thr Phe		
245	250	255
Asn Met His Gln Gln Ser Cys Lys Asp Leu Gln Leu Tyr Leu Ser Asn		
260	265	270
Leu Ala Asn Gln Ile Asp Arg Glu Thr Gly Ile Gly Asp Val Pro Leu		
275	280	285
Val Ile Leu Leu Asp Asp Leu Ser Glu Ala Gly Ser Ile Ser Glu Leu		
290	295	300
Val Asn Gly Ala Leu Thr Cys Lys Tyr His Lys Cys Pro Tyr Ile Ile		
305	310	315
320		
Gly Thr Thr Asn Gln Pro Val Lys Met Thr Pro Asn His Gly Leu His		
325	330	335
Leu Ser Phe Arg Met Leu Thr Phe Ser Asn Asn Val Glu Pro Ala Asn		
340	345	350
Gly Phe Leu Val Arg Tyr Leu Arg Arg Lys Leu Val Glu Ser Asp Ser		
355	360	365
Asp Ile Asn Ala Asn Lys Glu Glu Leu Leu Arg Val Leu Asp Trp Val		
370	375	380
Pro Lys Leu Trp Tyr His Leu His Thr Phe Leu Glu Lys His Ser Thr		
385	390	395
400		
Ser Asp Phe Leu Ile Gly Pro Cys Phe Phe Leu Ser Cys Pro Ile Gly		

405	410	415
Ile Glu Asp Phe Arg Thr Trp Phe Ile Asp Leu Trp Asn Asn Ser Ile		
420	425	430
Ile Pro Tyr Leu Gln Glu Gly Ala Lys Asp Gly Ile Lys Val His Gly		
435	440	445
Gln Lys Ala Ala Trp Glu Asp Pro Val Glu Trp Val Arg Asp Thr Leu		
450	455	460
Pro Trp Pro Ser Ala Gln Gln Asp Gln Ser Lys Leu Tyr His Leu Pro		
465	470	475
Pro Pro Thr Val Gly Pro His Ser Ile Ala Ser Pro Pro Glu Asp Arg		
485	490	495
Thr Val Lys Asp Ser Thr Pro Ser Ser Leu Asp Ser Asp Pro Leu Met		
500	505	510
Ala Met Leu Leu Lys Leu Gln Glu Ala Ala Asn Tyr Ile Glu Ser Pro		
515	520	525
Asp Arg Glu Thr Ile Leu Asp Pro Asn Leu Gln Ala Thr Leu		
530	535	540

<210> 7  
<211> 2568  
<212> DNA  
<213> mouse

<220>  
<221> CDS  
<223> (1)..(1695)

<400> 7

gaa cta tgg gaa aaa gag atg aag ctc acg gat atc cggttgg gag gcc	48
Glu Leu Trp Glu Lys Glu Met Lys Leu Thr Asp Ile Arg Leu Glu Ala	
ctc aac tct gcc cac cag ctg gac cag ctt cggttgg gag acc atg cac aat	96
Leu Asn Ser Ala His Gln Leu Asp Gln Leu Arg Glu Thr Met His Asn	
atg cag ttg gag gtg gac ctg ctg aaa gca gag aat gac cggttgg aag	144
Met Gln Leu Glu Val Asp Leu Leu Lys Ala Glu Asn Asp Arg Leu Lys	
gtt gcc ccc ggc ccc tcc tca ggc tgc act cca ggg cag gtc cct ggg	192
Val Ala Pro Gly Pro Ser Ser Gly Cys Thr Pro Gly Gln Val Pro Gly	

tca tcg gct ctg tcg tcc cct cga cgt tcc ctg ggc ctt gca ctc agc Ser Ser Ala Leu Ser Ser Pro Arg Arg Ser Leu Gly Leu Ala Leu Ser	240
cat cct ttc agt cct agt ctc aca gac aca gac ctc tca ccc atg gat His Pro Phe Ser Pro Ser Leu Thr Asp Thr Asp Leu Ser Pro Met Asp	288
ggc atc agc acc tgt ggt tca aag gaa gag gtg acc ctg cgg gtg gtg Gly Ile Ser Thr Cys Gly Ser Lys Glu Glu Val Thr Leu Arg Val Val	336
gtc cg <sup>g</sup> atg cc <sup>g</sup> ccc cag cac atc atc aaa gg <sup>g</sup> gac tta aag cag cag Val Arg Met Pro Pro Gln His Ile Lys Gly Asp Leu Lys Gln Gln	384
gag ttc ttc ctg ggt tgc agc aag gtc agt ggc aaa gtt gac tgg aag Glu Phe Phe Leu Gly Cys Ser Lys Val Ser Gly Lys Val Asp Trp Lys	432
atg ctg gat gaa gcc gtt ttc caa gtg ttc aag gac tac att tct aaa Met Leu Asp Glu Ala Val Phe Gln Val Phe Lys Asp Tyr Ile Ser Lys	480
atg gac cca gcc tca acc ctg gga ctg agc act gag tcc ata cat ggc Met Asp Pro Ala Ser Thr Leu Gly Leu Ser Thr Glu Ser Ile His Gly	528
tat agc ctc agc cac gtg aaa cga gtg ctg gat gct gag ccc cca gag Tyr Ser Leu Ser His Val Lys Arg Val Leu Asp Ala Glu Pro Pro Glu	576
atg cct cct tgc cgc cga ggt gtc aat aac ata tca gtc gct ctc aaa Met Pro Pro Cys Arg Arg Gly Val Asn Asn Ile Ser Val Ala Leu Lys	624
ggt ctg aaa gag aag tgt gtc gac agc ctg gtg ttc gag acg ctt atc Gly Leu Lys Glu Lys Cys Val Asp Ser Leu Val Phe Glu Thr Leu Ile	672
ccc aag ccc atg atg cag cac tac atc agc ctc ctg ctc aag cac cgg Pro Lys Pro Met Met Gln His Tyr Ile Ser Leu Leu Lys His Arg	720
cgc ctg gtg ctc tcc ggc ccc agt ggc acc ggc aag acc tac ttg acc Arg Leu Val Leu Ser Gly Pro Ser Gly Thr Gly Lys Thr Tyr Leu Thr	768
aat cgg cta gcc gag tac ctg gtg gag cgc tcc ggc cgc gag gtc acg Asn Arg Leu Ala Glu Tyr Leu Val Glu Arg Ser Gly Arg Glu Val Thr	816
gat ggc atc gtc agc act ttc aac atg cac cag cag tct tgc aag gat Asp Gly Ile Val Ser Thr Phe Asn Met His Gln Gln Ser Cys Lys Asp	864

ctg caa ctg tac ctc tcc aac cta gcc aac cag ata gac cg <sup>g</sup> gaa aca Leu Gln Leu Tyr Leu Ser Asn Leu Ala Asn Gln Ile Asp Arg Glu Thr	912
ggg ata ggg gat gtg ccc ttg gtg atc ctc ctg gat gat ctg agt gaa Gly Ile Gly Asp Val Pro Leu Val Ile Leu Leu Asp Asp Leu Ser Glu	960
gca ggc tcc atc agt gag ctg gtc aat ggg gcc ctc acc tgc aag tat Ala Gly Ser Ile Ser Glu Leu Val Asn Gly Ala Leu Thr Cys Lys Tyr	1008
cac aaa tgt ccc tac att ata ggt acc acc aat cag cct gta aaa atg His Lys Cys Pro Tyr Ile Ile Gly Thr Thr Asn Gln Pro Val Lys Met	1056
aca ccc aac cat ggc ttg cac ttg agc ttc agg atg ctg acc ttc tcg Thr Pro Asn His Gly Leu His Leu Ser Phe Arg Met Leu Thr Phe Ser	1104
aac aat gtg gaa cca gcc aat ggc ttt ctg gtc cgt tac ctg cgg agg Asn Asn Val Glu Pro Ala Asn Gly Phe Leu Val Arg Tyr Leu Arg Arg	1152
aag ttg gta gag tca gac agt gac gtc aat gct aac aag gaa gag ctg Lys Leu Val Glu Ser Asp Val Asn Ala Asn Lys Glu Glu Leu	1200
ctt cgg gtg ctg gac tgg gtg ccc aag ctg tgg tat cac ctc cac acc Leu Arg Val Leu Asp Trp Val Pro Lys Leu Trp Tyr His Leu His Thr	1248
ttc ctg gag aag cac acc tgc gac ttc ctc att ggc cct tgc ttc Phe Leu Glu Lys His Ser Thr Asp Phe Leu Ile Gly Pro Cys Phe	1296
ttc ctg tcc tgt ccc att ggc atc gag gac ttc cgg acc tgg ttc att Phe Leu Ser Cys Pro Ile Gly Ile Glu Asp Phe Arg Thr Trp Phe Ile	1344
gac ctg tgg aac aat tcc atc atc ccc tat cta cag gaa gga gcc aag Asp Leu Trp Asn Asn Ser Ile Ile Pro Tyr Leu Gln Glu Gly Ala Lys	1392
gat ggg atc aag gtt cat gga cag aaa gct gct tgg gaa gac ccg gtg Asp Gly Ile Lys Val His Gly Gln Lys Ala Ala Trp Glu Asp Pro Val	1440
gaa tgg gtc cga gac act ctt ccc tgg ccg tcg gcc caa caa gac caa Glu Trp Val Arg Asp Thr Leu Pro Trp Pro Ser Ala Gln Gln Asp Gln	1488
tca aag ctc tac cac ctg ccc ccg cct tct gtg ggc ccc cac agc act Ser Lys Leu Tyr His Leu Pro Pro Ser Val Gly Pro His Ser Thr	1536

gcc tca ccc ccg gag gac agg aca gtc aaa gac agc act cca aac tcc	1584
Ala Ser Pro Pro Glu Asp Arg Thr Val Lys Asp Ser Thr Pro Asn Ser	
ctc gac tca gat ccc ctg atg gcc atg cta ctg aaa ctc caa gaa gct	1632
Leu Asp Ser Asp Pro Leu Met Ala Met Leu Leu Lys Leu Gln Glu Ala	
gcc aac tac att gag tca cca gat cga gag act atc ctg gac ccc aac	1680
Ala Asn Tyr Ile Glu Ser Pro Asp Arg Glu Thr Ile Leu Asp Pro Asn	
ctc cag gcg aca ctc tgagggcccg gcagtcactg tcaccctgga gggcagaagg	1735
Leu Gln Ala Thr Leu	
ctggcttcag catcatttgc tctcctctgc cctcttcctt catagtcttg gtcaccagc	1795
ctcgccaaga gaacaggagg gaagaagagg gcaggaggag ggatgggttc tcggtgctga	1855
acctttgaga acttcctact aggaatttggg ggggtggag tttgagaact ccgtgcccct	1915
taactacatt tgctggcctc ctcttacgac ttaggagaaa agatgattct ggtctttct	1975
tcaagttttg tttcacctac aaactcttgg gctttctggg gagggattcg gaagatataa	2035
acagacaaac aaaaacaaac aaaccaaacta cagcagttcc aagctcgttc tcacaaacac	2095
ctctgagaca gtcacatgtg ggcaaattcta agggaggcag gaagctctac agactttctt	2155
gcaaaccctt cccagttctg tcgacactgc caacaacctc cccgccagag acctggccag	2215
agccaaagaaa agagaagcat gtggttAAC agaaaaacAA aacAAAacAA aacAAAAAAT	2275
atatgtgtAA atcaacctgt agaaggtaAA aacggcaatg gaaaagatga agctggagg	2335
aggggcccAG ttgccaAGAT ggaacgagAG ctgccaAGATC ttgccttCTG gatgacaAGA	2395
ggggacATTG caagatggCT gccagtctAA aacgtcacCA gaccacaAGA gtaacatCAC	2455
agccttcgAA gaaaggCCAC aagctgtCTT tctgccCTCT aactgaACAT gcatgAAAAG	2515
tcaataAAACC ctacttttAA attttAAaaa aaaaaaaaaa aaaaaaaaaa AAA	2568

<210> 8  
<211> 565  
<212> PRT  
<213> mouse

<400> 8

Glu	Leu	Trp	Glu	Lys	Glu	Met	Lys	Leu	Thr	Asp	Ile	Arg	Leu	Glu	Ala
1				5				10				15			

Leu	Asn	Ser	Ala	His	Gln	Leu	Asp	Gln	Leu	Arg	Glu	Thr	Met	His	Asn
						20		25				30			

Met	Gln	Leu	Glu	Val	Asp	Leu	Leu	Lys	Ala	Glu	Asn	Asp	Arg	Leu	Lys
35						40							45		
Val	Ala	Pro	Gly	Pro	Ser	Ser	Gly	Cys	Thr	Pro	Gly	Gln	Val	Pro	Gly
50						55							60		
Ser	Ser	Ala	Leu	Ser	Ser	Pro	Arg	Arg	Ser	Leu	Gly	Leu	Ala	Leu	Ser
65						70					75			80	
His	Pro	Phe	Ser	Pro	Ser	Leu	Thr	Asp	Thr	Asp	Leu	Ser	Pro	Met	Asp
						85				90				95	
Gly	Ile	Ser	Thr	Cys	Gly	Ser	Lys	Glu	Glu	Val	Thr	Leu	Arg	Val	Val
						100			105				110		
Val	Arg	Met	Pro	Pro	Gln	His	Ile	Ile	Lys	Gly	Asp	Leu	Lys	Gln	Gln
						115			120				125		
Glu	Phe	Phe	Leu	Gly	Cys	Ser	Lys	Val	Ser	Gly	Lys	Val	Asp	Trp	Lys
						130			135				140		
Met	Leu	Asp	Glu	Ala	Val	Phe	Gln	Val	Phe	Lys	Asp	Tyr	Ile	Ser	Lys
						145			150				155		160
Met	Asp	Pro	Ala	Ser	Thr	Leu	Gly	Leu	Ser	Thr	Glu	Ser	Ile	His	Gly
						165			170				175		
Tyr	Ser	Leu	Ser	His	Val	Lys	Arg	Val	Leu	Asp	Ala	Glu	Pro	Pro	Glu
						180			185				190		
Met	Pro	Pro	Cys	Arg	Arg	Gly	Val	Asn	Asn	Ile	Ser	Val	Ala	Leu	Lys
						195			200				205		
Gly	Leu	Lys	Glu	Lys	Cys	Val	Asp	Ser	Leu	Val	Phe	Glu	Thr	Leu	Ile
						210			215				220		
Pro	Lys	Pro	Met	Met	Gln	His	Tyr	Ile	Ser	Leu	Leu	Leu	Lys	His	Arg
						225			230				235		240
Arg	Leu	Val	Leu	Ser	Gly	Pro	Ser	Gly	Thr	Gly	Lys	Thr	Tyr	Leu	Thr
						245			250				255		
Asn	Arg	Leu	Ala	Glu	Tyr	Leu	Val	Glu	Arg	Ser	Gly	Arg	Glu	Val	Thr
						260			265				270		
Asp	Gly	Ile	Val	Ser	Thr	Phe	Asn	Met	His	Gln	Gln	Ser	Cys	Lys	Asp
						275			280				285		
Leu	Gln	Leu	Tyr	Leu	Ser	Asn	Leu	Ala	Asn	Gln	Ile	Asp	Arg	Glu	Thr
						290			295				300		
Gly	Ile	Gly	Asp	Val	Pro	Leu	Val	Ile	Leu	Leu	Asp	Asp	Leu	Ser	Glu
						305			310				315		320
Ala	Gly	Ser	Ile	Ser	Glu	Leu	Val	Asn	Gly	Ala	Leu	Thr	Cys	Lys	Tyr

325	330	335
His Lys Cys Pro Tyr Ile Ile Gly Thr Thr Asn Gln Pro Val Lys Met		
340	345	350
Thr Pro Asn His Gly Leu His Leu Ser Phe Arg Met Leu Thr Phe Ser		
355	360	365
Asn Asn Val Glu Pro Ala Asn Gly Phe Leu Val Arg Tyr Leu Arg Arg		
370	375	380
Lys Leu Val Glu Ser Asp Ser Asp Val Asn Ala Asn Lys Glu Glu Leu		
385	390	395
Leu Arg Val Leu Asp Trp Val Pro Lys Leu Trp Tyr His Leu His Thr		
405	410	415
Phe Leu Glu Lys His Ser Thr Ser Asp Phe Leu Ile Gly Pro Cys Phe		
420	425	430
Phe Leu Ser Cys Pro Ile Gly Ile Glu Asp Phe Arg Thr Trp Phe Ile		
435	440	445
Asp Leu Trp Asn Asn Ser Ile Ile Pro Tyr Leu Gln Glu Gly Ala Lys		
450	455	460
Asp Gly Ile Lys Val His Gly Gln Lys Ala Ala Trp Glu Asp Pro Val		
465	470	475
Glu Trp Val Arg Asp Thr Leu Pro Trp Pro Ser Ala Gln Gln Asp Gln		
485	490	495
Ser Lys Leu Tyr His Leu Pro Pro Pro Ser Val Gly Pro His Ser Thr		
500	505	510
Ala Ser Pro Pro Glu Asp Arg Thr Val Lys Asp Ser Thr Pro Asn Ser		
515	520	525
Leu Asp Ser Asp Pro Leu Met Ala Met Leu Leu Lys Leu Gln Glu Ala		
530	535	540
Ala Asn Tyr Ile Glu Ser Pro Asp Arg Glu Thr Ile Leu Asp Pro Asn		
545	550	555
Leu Gln Ala Thr Leu		
	565	

<210> 9  
 <211> 1025  
 <212> DNA  
 <213> mouse

<220>  
 <221> CDS  
 <223> (1)..(129)

<400> 9

cca ata gaa ctc cg <sup>g</sup> atc aag agg cag aat tcc tca gat agc atc tcc Pro Ile Glu Leu Arg Ile Lys Arg Gln Asn Ser Ser Asp Ser Ile Ser	48
agc ctc aac agc atc acc agc cat tcc agc atc ggc agc agc aaa gat Ser Leu Asn Ser Ile Thr Ser His Ser Ser Ile Gly Ser Ser Lys Asp	96
gct gat gcc aag aag aaa aag aag agt tgg gtatgtaaag gcttggggat Ala Asp Ala Lys Lys Lys Lys Lys Ser Trp	149
cggcctgtgc taggagtcac tcaccctgtt gcagggact gaccccttc aggatcaaca aagagggtcc cttctaacag gatgccagtg ttgtgacatc tgctggggac aaaaattcac	209
taagttccca ttccctatac cattgtctat tctccttacc accgcctgc acatataaccc cagcccccca ccgtccctgc atcccttata catgtctgct atccctgggc tctacctact	269
gatgaggtca aatgtat <sup>t</sup> tg gccgtagaag gagctgagaa aattattcat gggtgggaga gtggggcatg tggagagaat ttgtaagcca agcagggta <sup>c</sup> t ctagacgct cctggggctg	329
ttgctttagt ttgggtgagg aggctgtgga acgtccccat cgctccaaag cctgctttg tctggccag aggtgggtt gttctgtgt gtatcccccc tgtaactcta aactggctt	389
gggtgagctt tctacaatct gtacgcaggt gtagggca <sup>t</sup> t gcctgactga ctgaaaggga gagtgaccca gagtgagcgt cttgtccctg tccctgctga ggagggctgg ctacagactt	449
tggcctagtg cagacaggag ccagctgtgt ggagaagcag ctgtgtgaaa tgc <sup>t</sup> catgagta gtgtcgctgc tgctgctgct gctgcttct tttcattgtt ttttttttt tttcttcct	509
tttatttcct tcaaaatgct gaccta <sup>a</sup> at ccattttt tttccaggtt tatgaggtaa gaaactcgg <sup>t</sup> tcc <sup>t</sup> tcctc gtgctttc tttcccttg cacaccttcg tg <sup>t</sup> ttccaga	569
gcaagcacct ctctaaaaaa aaaaaaaaaa aaaaaaa	629
	689
	749
	809
	869
	929
	989
	1025

<210> 10

<211> 43  
<212> PRT  
<213> mouse

<400> 10

Pro Ile Glu Leu Arg Ile Lys Arg Gln Asn Ser Ser Asp Ser Ile Ser 1 5 10 15
Ser Leu Asn Ser Ile Thr Ser His Ser Ser Ile Gly Ser Ser Lys Asp

20

25

30

Ala Asp Ala Lys Lys Lys Lys Lys Ser Trp  
 35                          40

<210> 11  
 <211> 8690  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> CDS  
 <223> (503)..(6187)

<400> 11

ctctcagcgg cccaaggcagt ttctttctgg gtgacaagaa tgcgttcgg ttggttttc	60
tttttttctt ccatctcctt aagacgattt ccatacgtaac ctgatcaagt ggctcaaaat	120
cgc当地 cggaggccg cggcccggccg gcaagacctc ggccaggttt acaggaaatc	180
tgtcattttt tattaaaatg gaaaactgtg aagaaagaaa aagatagcag ttgaagtcaa	240
aattctcgga tgactatTTT gctttgagg agtcagcatt taAAAacgat atgctgattt	300
ggaaggcctt gggagtaaac tgcaaacttt atttttcca ttcaatcaat ggattttta	360
atcattcctt ggagtcgatg aagttcgaa acgggtgtgtg atggggaaacg tggcgggcca	420
gtgtgttcctt agaaatttgcg tcttgattt gtttgctgct ttttgaaga gattccattt	480
tgaaggcggaa gaacctaatg tg atg gat tta tct tca gaa atg aac aga cat	532
Met Asp Leu Ser Ser Glu Met Asn Arg His	
ggg aag aat cca gtg agt cac aag cta gaa gat cag aag aag att tac	580
Gly Lys Asn Pro Val Ser His Lys Leu Glu Asp Gln Lys Lys Ile Tyr	
act gac tgg gcc aac cac tac cta gca aaa tca ggc cac aag cgg ctg	628
Thr Asp Trp Ala Asn His Tyr Leu Ala Lys Ser Gly His Lys Arg Leu	
atc aag gac ttg caa caa gac att gca gat gga gta ctc cta gca gaa	676
Ile Lys Asp Leu Gln Gln Asp Ile Ala Asp Gly Val Leu Leu Ala Glu	
atc atc cag att att gca aat gaa aaa gtt gaa gat atc aat gga tgt	724
Ile Ile Gln Ile Ile Ala Asn Glu Lys Val Glu Asp Ile Asn Gly Cys	
cct aga agt cag tct cag atg att gaa aat gtt gat gtc tgc ctt agt	772
Pro Arg Ser Gln Ser Gln Met Ile Glu Asn Val Asp Val Cys Leu Ser	

ttt cta gca gcc aga ggg gta aat gtt caa ggt cta tct gct gaa gaa	820
Phe Leu Ala Ala Arg Gly Val Asn Val Gln Gly Leu Ser Ala Glu Glu	
ata aga aat gga aac tta aaa gcc att cta ggg ctg ttt ttc agt tta	868
Ile Arg Asn Gly Asn Leu Lys Ala Ile Leu Gly Leu Phe Phe Ser Leu	
tct cgc tac aag cag caa caa cac cat caa caa cag tac tat cag tcc	916
Ser Arg Tyr Lys Gln Gln His His Gln Gln Tyr Tyr Gln Ser	
ttg gtg gaa ctt cag cag cga gtt act cac gct tcc cct cca tcg gaa	964
Leu Val Glu Leu Gln Gln Arg Val Thr His Ala Ser Pro Pro Ser Glu	
gcc agc cag gcc aaa acc cag caa gat atg cag tcc agg ctt cca ggg	1012
Ala Ser Gln Ala Lys Thr Gln Gln Asp Met Gln Ser Arg Leu Pro Gly	
ccc tct agg gtg cct gct gca gga agc agc aag gtc cag gga gcc	1060
Pro Ser Arg Val Pro Ala Ala Gly Ser Ser Lys Val Gln Gly Ala	
tct aat tta aat agg aga agt cag agc ttt aac agc att gac aaa aac	1108
Ser Asn Leu Asn Arg Ser Gln Ser Phe Asn Ser Ile Asp Lys Asn	
aag cct cca aat tat gca aat gga aac gaa aaa ggt gaa gac cct gaa	1156
Lys Pro Pro Asn Tyr Ala Asn Gly Asn Glu Lys Gly Glu Asp Pro Glu	
aca aga aga atg aga aca gtt aaa aac ata gca gac ttg agg cag aat	1204
Thr Arg Arg Met Arg Thr Val Lys Asn Ile Ala Asp Leu Arg Gln Asn	
tta gaa gag act atg tcc agt ctt cgt ggg act cag ata agc cac agc	1252
Leu Glu Glu Thr Met Ser Ser Leu Arg Gly Thr Gln Ile Ser His Ser	
acc ctg gag aca aca ttt gac agc act gtg aca aca gaa gtt aat gga	1300
Thr Leu Glu Thr Phe Asp Ser Thr Val Thr Glu Val Asn Gly	
agg acc ata ccc aac ttg aca agt cga ccc acc ccc atg acc tgg agg	1348
Arg Thr Ile Pro Asn Leu Thr Ser Arg Pro Thr Pro Met Thr Trp Arg	
ttg ggc cag gca tgt ccg cga ctt cag gcg gga gat gct ccc tcc ctg	1396
Leu Gly Gln Ala Cys Pro Arg Leu Gln Ala Gly Asp Ala Pro Ser Leu	
ggt gct ggc tat cct cgc agt ggt acc agt cga ttc atc cac aca gac	1444
Gly Ala Gly Tyr Pro Arg Ser Gly Thr Ser Arg Phe Ile His Thr Asp	

ccc tcg agg ttc atg tat acc acg cct ctc cgt cga gct gct gtc tct	1492
Pro Ser Arg Phe Met Tyr Thr Pro Leu Arg Arg Ala Ala Val Ser	
agg ctg gga aac atg tca cag att gac atg agt gag aaa gca agc agt	1540
Arg Leu Gly Asn Met Ser Gln Ile Asp Met Ser Glu Lys Ala Ser Ser	
gac ctg gac atg tct tct gag gtc gat gtg ggt gga tat atg agt gat	1588
Asp Leu Asp Met Ser Ser Glu Val Asp Val Gly Gly Tyr Met Ser Asp	
ggt gat atc ctt ggg aaa agt ctc agg act gat gac atc aac agt ggg	1636
Gly Asp Ile Leu Gly Lys Ser Leu Arg Thr Asp Asp Ile Asn Ser Gly	
tac atg aca gat gga gga ctt aac cta tat act aga agt ctg aac cga	1684
Tyr Met Thr Asp Gly Leu Asn Leu Tyr Thr Arg Ser Leu Asn Arg	
ata cca gac aca gca act tcc cgg gac atc atc cag aga ggg gtt cac	1732
Ile Pro Asp Thr Ala Thr Ser Arg Asp Ile Ile Gln Arg Gly Val His	
gat gtg aca gtg gat gca gac agc tgg gat gac agc agt tca gtg agc	1780
Asp Val Thr Val Asp Ala Asp Ser Trp Asp Asp Ser Ser Val Ser	
agt ggt ctc agt gac acc ctt gat aac atc agc act gat gac ctg aac	1828
Ser Gly Leu Ser Asp Thr Leu Asp Asn Ile Ser Thr Asp Asp Leu Asn	
acc aca tcc tct gtc agc tct tac tcc aac atc acc gtc ccc tct agg	1876
Thr Thr Ser Ser Val Ser Tyr Ser Asn Ile Thr Val Pro Ser Arg	
aag aat act cag ctg agg aca gat tca gag aaa cgc tcc acc aca gac	1924
Lys Asn Thr Gln Leu Arg Thr Asp Ser Glu Lys Arg Ser Thr Thr Asp	
gag acc tgg gat agt cct gag gaa ctg aaa aaa cca gaa gaa gat ttt	1972
Glu Thr Trp Asp Ser Pro Glu Glu Leu Lys Lys Pro Glu Glu Asp Phe	
gac agc cat ggg gat gct ggt ggc aag tgg aag act gtg tcc tct gga	2020
Asp Ser His Gly Asp Ala Gly Lys Trp Lys Thr Val Ser Ser Gly	
ctt cct gaa gac ccc gag aag gca ggg cag aaa gct tcc ctg tct gtt	2068
Leu Pro Glu Asp Pro Glu Lys Ala Gly Gln Lys Ala Ser Leu Ser Val	
tca cag aca ggt tcc tgg aga aga ggc atg tct gcc caa gga ggg gcg	2116
Ser Gln Thr Gly Ser Trp Arg Arg Gly Met Ser Ala Gln Gly Gly Ala	

cca tct agg cag aaa gct gga aca agt gca ctc aaa aca ccc ggg aaa Pro Ser Arg Gln Lys Ala Gly Thr Ser Ala Leu Lys Thr Pro Gly Lys	2164
acc gat gat gcc aaa gct tct gag aaa gga aaa gct ccc cta aaa gga Thr Asp Asp Ala Lys Ala Ser Glu Lys Gly Ala Pro Leu Lys Gly	2212
tca tct cta caa aga tct cct tca gat gca gga aaa agc agt gga gat Ser Ser Leu Gln Arg Ser Pro Asp Ala Gly Lys Ser Ser Gly Asp	2260
gaa ggg aaa aag ccc ccc tca ggc att gga aga tcg act gcc acc agc Glu Gly Lys Lys Pro Pro Ser Gly Ile Gly Arg Ser Thr Ala Thr Ser	2308
tcc ttt ggc ttt aag aaa cca agt gga gta ggg tca tct gcc atg atc Ser Phe Gly Phe Lys Lys Pro Ser Gly Val Gly Ser Ser Ala Met Ile	2356
acc agc agt gga gca acc ata aca agt ggc tct gca aca ctg ggt aaa Thr Ser Ser Gly Ala Thr Ile Thr Ser Gly Ser Ala Thr Leu Gly Lys	2404
att cca aaa tct gct gcc att ggc ggg aag tca aat gca ggg aga aaa Ile Pro Lys Ser Ala Ala Ile Gly Lys Ser Asn Ala Gly Arg Lys	2452
acc agt ttg gac ggt tca cag aat cag gat gat gtt gtg ctg cat gtt Thr Ser Leu Asp Gly Ser Gln Asn Gln Asp Asp Val Val Leu His Val	2500
agc tca aag act acc cta caa tat cgc agc ttg ccc cgc cct tca aaa Ser Ser Lys Thr Thr Leu Gln Tyr Arg Ser Leu Pro Arg Pro Ser Lys	2548
tcc agc acc agt ggc att cct ggc cga gga ggc cac aga tcc agt acc Ser Ser Thr Ser Gly Ile Pro Gly Arg Gly His Arg Ser Ser Thr	2596
agc agt att gat tcc aac gtc agc agc aag tct gct ggg gcc acc acc Ser Ser Ile Asp Ser Asn Val Ser Ser Lys Ser Ala Gly Ala Thr Thr	2644
tcg aaa ctg aga gaa cca act aaa att ggg tca ggg cgc tcg agt cct Ser Lys Leu Arg Glu Pro Thr Lys Ile Gly Ser Gly Arg Ser Ser Pro	2692
gtc acc gtc aac caa aca gac aag gaa aag gaa aaa gta gca gtc tca Val Thr Val Asn Gln Thr Asp Lys Glu Lys Val Ala Val Ser	2740
gat tca gaa agt gtt tct ttg tca ggt tcc ccc aaa tcc agc ccc acc Asp Ser Glu Ser Val Ser Leu Ser Gly Ser Pro Lys Ser Ser Pro Thr	2788

tct gcc agc gcc tgt ggt gca caa ggt ctc agg cag cca gga tcc aag Ser Ala Ser Ala Cys Gly Ala Gln Gly Leu Arg Gln Pro Gly Ser Lys	2836
tat cca gat att gcc tca ccc aca ttt cga agg ttg ttt ggt gcc aag Tyr Pro Asp Ile Ala Ser Pro Thr Phe Arg Arg Leu Phe Gly Ala Lys	2884
gca ggt ggc aaa tct gcc tct gca cct aat act gag ggt gtg aaa tct Ala Gly Gly Lys Ser Ala Ser Ala Pro Asn Thr Glu Gly Val Lys Ser	2932
tcc tca gta atg ccc agc cct agt acc aca tta gcg cggt caa ggc agt Ser Ser Val Met Pro Ser Pro Ser Thr Thr Leu Ala Arg Gln Gly Ser	2980
ctg gag tca ccg tcg tcc ggt acg ggc agc atg ggc agt gct ggt ggg Leu Glu Ser Pro Ser Ser Gly Thr Gly Ser Met Gly Ser Ala Gly Gly	3028
cta agc ggc agc agc cct ctc ttc aat aaa ccc tca gac tta act Leu Ser Gly Ser Ser Pro Leu Phe Asn Lys Pro Ser Asp Leu Thr	3076
aca gat gtt ata agc tta agt cac tcg ttg gcc tcc agc cca gca tcg Thr Asp Val Ile Ser Leu Ser His Ser Leu Ala Ser Ser Pro Ala Ser	3124
gtt cac tct ttc aca tca ggt ggt ctc gtg tgg gct gcc aat atg agc Val His Ser Phe Thr Ser Gly Gly Leu Val Trp Ala Ala Asn Met Ser	3172
agt tcc tct gca ggc agc aag gat act ccg agc tac cag tcc atg act Ser Ser Ser Ala Gly Ser Lys Asp Thr Pro Ser Tyr Gln Ser Met Thr	3220
agc ctc cac acg agc tct gag tcc att gac ctc ccc ctc agc cat cat Ser Leu His Thr Ser Ser Glu Ser Ile Asp Leu Pro Leu Ser His His	3268
ggc tcc ttg tct gga ctg acc aca ggc act cac gag gtc cag agc ctg Gly Ser Leu Ser Gly Leu Thr Thr Gly Thr His Glu Val Gln Ser Leu	3316
ctc atg aga acg ggt agt gtg aga tct act ctc tca gaa agc atg cag Leu Met Arg Thr Gly Ser Val Arg Ser Thr Leu Ser Glu Ser Met Gln	3364
ctt gac aga aat aca cta ccc aaa aag gga cta aga tat acc cca tca Leu Asp Arg Asn Thr Leu Pro Lys Lys Gly Leu Arg Tyr Thr Pro Ser	3412
tct cgg cag gcc aac caa gaa gag ggc aaa gag tgg ttg cgt tct cat Ser Arg Gln Ala Asn Gln Glu Glu Gly Lys Glu Trp Leu Arg Ser His	3460
tct act gga ggg ctt cag gac act ggc aac cag tca cct ctg gtt tcc	3508

Ser Thr Gly Gly Leu Gln Asp Thr Gly Asn Gln Ser Pro Leu Val Ser	
cct tct gcc atg tca tct tct gca gct gga aaa tac cac ttt tct aac Pro Ser Ala Met Ser Ser Ala Ala Gly Lys Tyr His Phe Ser Asn	3556
ttg gtg agc cca aca aat ttg tct caa ttt aac ctt ccc ggg ccc agc Leu Val Ser Pro Thr Asn Leu Ser Gln Phe Asn Leu Pro Gly Pro Ser	3604
atg atg cgc tca aac agc atc cca gcc caa gac tct tcc ttc gat ctc Met Met Arg Ser Asn Ser Ile Pro Ala Gln Asp Ser Phe Asp Leu	3652
tat gat gac tcc cag ctt tgt ggg agt gcc act tct ctg gag gaa aga Tyr Asp Asp Ser Gln Leu Cys Gly Ser Ala Thr Ser Leu Glu Glu Arg	3700
cct cgt gcc atc agt cat tcg ggc tca ttc aga gac agc atg gaa gaa Pro Arg Ala Ile Ser His Ser Gly Ser Phe Arg Asp Ser Met Glu Glu	3748
gtt cat ggc tct tca tta tca ctg gtg tcc agc act tct tct ctt tac Val His Gly Ser Ser Leu Ser Leu Val Ser Ser Thr Ser Ser Leu Tyr	3796
tct aca gct gaa gaa aag gct cat tca gag caa atc cat aaa ctg cgg Ser Thr Ala Glu Glu Lys Ala His Ser Glu Gln Ile His Lys Leu Arg	3844
aga gag ctg gtt gca tca caa gaa aaa gtt gct acc ctc aca tct cag Arg Glu Leu Val Ala Ser Gln Glu Lys Val Ala Thr Leu Thr Ser Gln	3892
ctt tca gca aat gct cac ctt gta gca gct ttt gaa aag agc tta ggg Leu Ser Ala Asn Ala His Leu Val Ala Ala Phe Glu Lys Ser Leu Gly	3940
aat atg act ggc cga ttg caa agt cta act atg aca gcg gaa caa aag Asn Met Thr Gly Arg Leu Gln Ser Leu Thr Met Thr Ala Glu Gln Lys	3988
gaa tct gaa ctt ata gaa cta aga gaa acc att gaa atg ctg aag gct Glu Ser Glu Leu Ile Glu Leu Arg Glu Thr Ile Glu Met Leu Lys Ala	4036
cag aat tct gct gcc cag gcg gct att cag gga gca ctg aat ggt cca Gln Asn Ser Ala Ala Gln Ala Ile Gln Gly Ala Leu Asn Gly Pro	4084
gac cat cct ccc aaa gat ctt cgc atc aga aga cag cat tcc tct gaa Asp His Pro Pro Lys Asp Leu Arg Ile Arg Arg Gln His Ser Ser Glu	4132
agt gtt tct agt atc aac agt gcc aca agc cat tcc agt att ggc agt Ser Val Ser Ser Ile Asn Ser Ala Thr Ser His Ser Ser Ile Gly Ser	4180

ggt aat gat gcc gac tcc aag aag aag aaa aac aac tgg ctg aga	4228
Gly Asn Asp Ala Asp Ser Lys Lys Lys Lys Lys Asn Trp Leu Arg	
agt tct ttc aaa caa gcc ttt ggg aag aaa aag tcc acc aag cct cct	4276
Ser Ser Phe Lys Gln Ala Phe Gly Lys Lys Ser Thr Lys Pro Pro	
tca tca cat tct gac att gaa gag ctt act gat tca tcc ctt ccg gca	4324
Ser Ser His Ser Asp Ile Glu Glu Leu Thr Asp Ser Ser Leu Pro Ala	
tcc ccc aag tta ccc cat aat gct ggt gac tgt ggc tca gca tcc atg	4372
Ser Pro Lys Leu Pro His Asn Ala Gly Asp Cys Gly Ser Ala Ser Met	
aag ccc tca caa tct gct tca gcg atc tgt gaa tgc aca gaa gct gag	4420
Lys Pro Ser Gln Ser Ala Ser Ala Ile Cys Glu Cys Thr Glu Ala Glu	
gca gag ata att ctg cag ctg aag agc gag ctc aga gaa aag gaa tta	4468
Ala Glu Ile Ile Leu Gln Leu Lys Ser Glu Leu Arg Glu Lys Glu Leu	
aaa tta acg gat att cgg ctg gag gcc ctc agc tct gct cat cat ctt	4516
Lys Leu Thr Asp Ile Arg Leu Glu Ala Leu Ser Ser Ala His His Leu	
gat cag atc cgg gaa gcc atg aac cgg atg cag aat gaa att gaa ata	4564
Asp Gln Ile Arg Glu Ala Met Asn Arg Met Gln Asn Glu Ile Glu Ile	
ctg aaa gct gaa aat gac cgg ttg aag gca gaa act ggt aac aca gct	4612
Leu Lys Ala Glu Asn Asp Arg Leu Lys Ala Glu Thr Gly Asn Thr Ala	
aag cct act cgg cca ccg tca gaa tcc tca agc agc acc tcc tct tca	4660
Lys Pro Thr Arg Pro Pro Ser Glu Ser Ser Ser Thr Ser Ser Ser	
tct tcc agg cag tca tta gga ctt tct cta aac aat ttg aac atc aca	4708
Ser Ser Arg Gln Ser Leu Gly Leu Ser Leu Asn Asn Leu Asn Ile Thr	
gag gct gtt agc tca gat att ttg cta gat gat gct ggt gat gca act	4756
Glu Ala Val Ser Ser Asp Ile Leu Leu Asp Asp Ala Gly Asp Ala Thr	
gga cat aaa gat ggc cgc agt gtg aaa att ata gtc tcc ata agc aag	4804
Gly His Lys Asp Gly Arg Ser Val Lys Ile Ile Val Ser Ile Ser Lys	
ggc tat ggt cga gca aag gac caa aaa tct cag gca tat ttg ata gga	4852
Gly Tyr Gly Arg Ala Lys Asp Gln Lys Ser Gln Ala Tyr Leu Ile Gly	

tcc att ggt gtt agt gga aaa acc aag tgg gat gtc tta gat ggt gta	4900
Ser Ile Gly Val Ser Gly Lys Thr Lys Trp Asp Val Leu Asp Gly Val	
ata aga cgt ctc ttt aag gaa tat gta ttc cga att gat aca tcc act	4948
Ile Arg Arg Leu Phe Lys Glu Tyr Val Phe Arg Ile Asp Thr Ser Thr	
agc ctt ggt ctg agc tct gac tgc att gct agc tac tgt ata gga gac	4996
Ser Leu Gly Leu Ser Ser Asp Cys Ile Ala Ser Tyr Cys Ile Gly Asp	
tta att aga tcc cat aac cta gaa gtg cct gaa ttg ctg cct tgt gga	5044
Leu Ile Arg Ser His Asn Leu Glu Val Pro Glu Leu Leu Pro Cys Gly	
tac ctt gtt gga gat aat aac atc atc act gtg aac ctc aaa ggg gta	5092
Tyr Leu Val Gly Asp Asn Ile Ile Thr Val Asn Leu Lys Gly Val	
gaa gaa aat agt ttg gac agt ttt gtt ttt gat acg ctg att cct aaa	5140
Glu Glu Asn Ser Leu Asp Ser Phe Val Phe Asp Thr Leu Ile Pro Lys	
cca att acc caa agg tac ttt aac ttg ttg atg gag cat cac aga att	5188
Pro Ile Thr Gln Arg Tyr Phe Asn Leu Leu Met Glu His His Arg Ile	
ata ctc tca gga ccg agt ggt act gga aag acc tat ttg gca aac aaa	5236
Ile Leu Ser Gly Pro Ser Gly Thr Tyr Leu Ala Asn Lys	
ctt gct gaa tat gta ata acc aaa tct gga agg aaa aaa aca gag gat	5284
Leu Ala Glu Tyr Val Ile Thr Lys Ser Gly Arg Lys Lys Thr Glu Asp	
gca att gcc act ttt aat gtg gac cac aag tca agt aag gaa ttg caa	5332
Ala Ile Ala Thr Phe Asn Val Asp His Lys Ser Ser Lys Glu Leu Gln	
caa tat cta gct aac ctg gct gaa cag tgc agt gct gat aat aat gga	5380
Gln Tyr Leu Ala Asn Leu Ala Glu Gln Cys Ser Ala Asp Asn Asn Gly	
gtg gag ctc cca gtt gta ata att ctt gat aat ctt cat cat gtg ggc	5428
Val Glu Leu Pro Val Val Ile Ile Leu Asp Asn Leu His His Val Gly	
tct ctg agt gat atc ttc aat ggt ttt ctc aat tgt aaa tac aac aaa	5476
Ser Leu Ser Asp Ile Phe Asn Gly Phe Leu Asn Cys Lys Tyr Asn Lys	
tgt cca tat att att gga aca atg aat cag gga gtt tct tca tca cca	5524
Cys Pro Tyr Ile Ile Gly Thr Met Asn Gln Gly Val Ser Ser Pro	

aat cta gag ctg cat cac aat ttc agg tgg gta tta tgt gca aat cat Asn Leu Glu Leu His His Asn Phe Arg Trp Val Leu Cys Ala Asn His	5572
aca gaa cca gtg aaa ggc ttt tta ggc aga tat ctt cga aga aaa ctc Thr Glu Pro Val Lys Gly Phe Leu Gly Arg Tyr Leu Arg Arg Lys Leu	5620
ata gag ata gaa att gaa agg aac att cgc aat aat gac cta gtc aaa Ile Glu Ile Glu Ile Glu Arg Asn Ile Arg Asn Asn Asp Leu Val Lys	5668
att ata gat tgg att ccg aag acg tgg cat cat ctc aac agt ttt ttg Ile Ile Asp Trp Ile Pro Lys Thr Trp His His Leu Asn Ser Phe Leu	5716
gaa aca cac agt tct tct gac gtt acc att ggt ccc cga cta ttc ctt Glu Thr His Ser Ser Asp Val Thr Ile Gly Pro Arg Leu Phe Leu	5764
cct tgc ccc atg gat gta gaa ggt tct aga gta tgg ttc atg gat ctc Pro Cys Pro Met Asp Val Glu Gly Ser Arg Val Trp Phe Met Asp Leu	5812
tgg aac tat tct tta gta cct tat att ctg gag gca gtg aga gag ggt Trp Asn Tyr Ser Leu Val Pro Tyr Ile Leu Glu Ala Val Arg Glu Gly	5860
ctt cag atg tat ggg aaa cgc aca cca tgg gaa gat cct tca aag tgg Leu Gln Met Tyr Gly Lys Arg Thr Pro Trp Glu Asp Pro Ser Lys Trp	5908
gtg ctt gac aca tat cca tgg agc tca gca act ctg cct cag gag agc Val Leu Asp Thr Tyr Pro Trp Ser Ser Ala Thr Leu Pro Gln Glu Ser	5956
cca gcc tta ctt cag ctg cga cca gaa gat gtt ggg tat gaa agc tgc Pro Ala Leu Leu Gln Leu Arg Pro Glu Asp Val Gly Ser Cys	6004
aca tcc act aag gaa gcc aca acc tca aag cac att cca caa act gac Thr Ser Thr Lys Glu Ala Thr Thr Ser Lys His Ile Pro Gln Thr Asp	6052
aca gaa gga gat ccc ctg atg aat atg cta atg aaa ctc caa gaa gca Thr Glu Gly Asp Pro Leu Met Asn Met Leu Met Lys Leu Gln Glu Ala	6100
gcc aat tac tcg agc aca caa agc tgc gac agc gaa agc acc agt cac Ala Asn Tyr Ser Ser Thr Gln Ser Cys Asp Ser Glu Ser Thr Ser His	6148
cat gaa gac att ttg gat tca tct ctt gaa tct acc ctc tagagggtga His Glu Asp Ile Leu Asp Ser Ser Leu Glu Ser Thr Leu	6197
aaaaagttaa gggaaaagac ttgcgtttta aaaaaatgtt tcaaaagaaa ggtatTTCA	6257

ctaaaccact gccagtataa aagcacccctg tcaaggggccc tgacccagag ttgtggtctc	6317
caaggaggca gcagaactaa gtctgaaccg ccaagatgct aaattgcaat ggaagcttaa	6377
ctttagtttta tttctaaaca ttttttatat ctgtggagta atagaaaagct ccattactca	6437
actggaaagg accctaattga cagggcaact gaacagattg cacatggat agccaaactg	6497
gactttcttt gtttcctctt taaaagtttta caatgcagac catttttgt cccttccttt	6557
tgtttcctct gaggggctgt tcgccccagg cagggccat ctttctgatc tgtccaacct	6617
cctttgtgcc acacggtgct ggtcacaggg cttcagtagt gtttgtgtg tgcgctcacc	6677
ccattccaga acaaattccaa gaggccagtc ctccataagc acaaattggaa ttgtcaacc	6737
accagaaaaa cactactgtg gcaaactggaa gaagtgc当地 ttaattcta actgccacgt	6797
tctcatgatg tgctccacca acttttttagt atatgagtca ctggtttat aaggttgtt	6857
ttaccacagt ggtctttta aaccacctgc ccactccctt aacaagagtt ttataccaat	6917
tattagtcaa cactgataaa aggtttttt agggctttat ttgtttgagc ctttcagtg	6977
aaagaaggaa catttcctat ggtgctgtct cactgcctta aaacagattt ctatgacagt	7037
ttaacagttg gtttaaatcc taaaccattt gtaatttcca ctgtctttc atttacaacc	7097
aagcaacacc agttaacata gtagcctcat ctctatataat ctttctctt ttttttttt	7157
tgaagaaatg gataggagaa agatcagttt ttttagcctt gtgaatagat cgcttgcct	7217
atcctccaaa atattaaaat aacccagaaaa tgctcttga ccgtcactta aaacctaaga	7277
catgtggcga aattccatcc agttctaagt gaaagagttt cagaaggcag gagattttga	7337
attattatcc agcagggctg gaagcactag atgcagcatg agcacaacta ttccggcttc	7397
ctccctatt gttttgttt ttttaatgag ttttgacgca tttttttt attgctattt	7457
ttgtacatga gaaattcagc attaaagaac actgaagcgg taaggtcact gtggaaagagg	7517
aagcgtttat actgtaaaag aaggtagat ttgcacagtc tactggtag gtattgtaaa	7577
taataatttt taaaacttgc acaaattccaa acaaacacaa acaaattgt attttatcct	7637
gttgggttta agaggtttt cacttgctga gatttcctgt acattgc当地 caaatacaga	7697
atgcaaacc tcaaagctgt attatctggt gtgtttgtcc ttttataca gttttttt	7757
actatgcagg agctatcagt gctagagtga gcatgcttca aaactgtaca tgaagccaat	7817
atattttgg ataagtaaaa ctgtctgaaa gtacatctgt catggcaggc tttaaagaga	7877
gtgcatgaaa actgatcagt cattggagaa gttaccacca cacacaaagg acaggttttta	7937

agttttatgaa acccaagggc taggccatgg tatagacttc ttctatgagt gtgtgaaaat	7997
gtgttacttt taggacgtgt atttggtgct actctctgtg accaccaatg gtcagttgc	8057
tatagaacaa caacaccacg aaacatctgt gcagtttca gagtgcaca aagtcaatag	8117
gtccttacac ggtgctattt ccctaaggga aatccgaact gaatttatgc acatagaatt .	8177
gtcacccctga ctttgaagcc tcaaacatgg atcaaatctg ttgtgaaaca tcaatatatg	8237
tagctggatg agtgactagt ttcccttgta taatatgtga tctaagaaaa ttgctaatct	8297
ttccctgcca ttttgagaaa cacagtccaa acatgagcat aaacagaatt tcctgcaata	8357
catcccagta ggtccaccta gtttacaact taaacttagtt tgtgaaacat ttgtctgtat	8417
acattttata ttttgtacat tttgatgtaa catatcatgt aaataggcag aaacagtcaa	8477
ataaaatcatc tgaaaagttt tgttagtcttt gtaaagcccc aacaataagt acttgggtgc	8537
aatggactta actggatgat gtatttcta ttggtttatt gttcctctag cttgtaaacc	8597
agcttgcata tatttttttg caaatgtgca ccctgtatct gtctaaatta ttactttgcc	8657
attaaaqtqg aatttatttat tgacaaaaaaaaaaa	8690

<210> 12  
<211> 1895  
<212> PRT  
<213> *Homo sapiens*

<400> 12

Met Asp Leu Ser Ser Glu Met Asn Arg His Gly Lys Asn Pro Val Ser  
1 5 10 15

His Lys Leu Glu Asp Gln Lys Lys Ile Tyr Thr Asp Trp Ala Asn His  
20 25 30

Tyr Leu Ala Lys Ser Gly His Lys Arg Leu Ile Lys Asp Leu Gln Gln  
 35 40 45

Asp Ile Ala Asp Gly Val Leu Leu Ala Glu Ile Ile Gln Ile Ile Ala  
50 55 60

Asn	Glu	Lys	Val	Glu	Asp	Ile	Asn	Gly	Cys	Pro	Arg	Ser	Gln	Ser	Gln
65				70						75					80

Met Ile Glu Asn Val Asp Val Cys Leu Ser Phe Leu Ala Ala Arg Gly  
85 90 95

Val Asn Val Gln Gly Leu Ser Ala Glu Glu Ile Arg Asn Gly Asn Leu  
100 105 110

Lys Ala Ile Leu Gly Leu Phe Phe Ser Leu Ser Arg Tyr Lys Gln Gln

115	120	125
Gln His His Gln Gln	Tyr Tyr Gln Ser Leu Val Glu Leu Gln Gln	
130	135	140
Arg Val Thr His Ala Ser Pro Pro Ser Glu Ala Ser Gln Ala Lys Thr		
145	150	160
Gln Gln Asp Met Gln Ser Arg Leu Pro Gly Pro Ser Arg Val Pro Ala		
165	170	175
Ala Gly Ser Ser Ser Lys Val Gln Gly Ala Ser Asn Leu Asn Arg Arg		
180	185	190
Ser Gln Ser Phe Asn Ser Ile Asp Lys Asn Lys Pro Pro Asn Tyr Ala		
195	200	205
Asn Gly Asn Glu Lys Gly Glu Asp Pro Glu Thr Arg Arg Met Arg Thr		
210	215	220
Val Lys Asn Ile Ala Asp Leu Arg Gln Asn Leu Glu Glu Thr Met Ser		
225	230	240
Ser Leu Arg Gly Thr Gln Ile Ser His Ser Thr Leu Glu Thr Thr Phe		
245	250	255
Asp Ser Thr Val Thr Thr Glu Val Asn Gly Arg Thr Ile Pro Asn Leu		
260	265	270
Thr Ser Arg Pro Thr Pro Met Thr Trp Arg Leu Gly Gln Ala Cys Pro		
275	280	285
Arg Leu Gln Ala Gly Asp Ala Pro Ser Leu Gly Ala Gly Tyr Pro Arg		
290	295	300
Ser Gly Thr Ser Arg Phe Ile His Thr Asp Pro Ser Arg Phe Met Tyr		
305	310	320
Thr Thr Pro Leu Arg Arg Ala Ala Val Ser Arg Leu Gly Asn Met Ser		
325	330	335
Gln Ile Asp Met Ser Glu Lys Ala Ser Ser Asp Leu Asp Met Ser Ser		
340	345	350
Glu Val Asp Val Gly Gly Tyr Met Ser Asp Gly Asp Ile Leu Gly Lys		
355	360	365
Ser Leu Arg Thr Asp Asp Ile Asn Ser Gly Tyr Met Thr Asp Gly Gly		
370	375	380
Leu Asn Leu Tyr Thr Arg Ser Leu Asn Arg Ile Pro Asp Thr Ala Thr		
385	390	400
Ser Arg Asp Ile Ile Gln Arg Gly Val His Asp Val Thr Val Asp Ala		
405	410	415
Asp Ser Trp Asp Asp Ser Ser Ser Val Ser Ser Gly Leu Ser Asp Thr		

420	425	430
Leu Asp Asn Ile Ser Thr Asp Asp	Leu Asn Thr Thr Ser Ser Val Ser	
435	440	445
Ser Tyr Ser Asn Ile Thr Val Pro Ser Arg Lys Asn Thr Gln Leu Arg		
450	455	460
Thr Asp Ser Glu Lys Arg Ser Thr Thr Asp Glu Thr Trp Asp Ser Pro		
465	470	475
Glu Glu Leu Lys Lys Pro Glu Glu Asp Phe Asp Ser His Gly Asp Ala		
485	490	495
Gly Gly Lys Trp Lys Thr Val Ser Ser Gly Leu Pro Glu Asp Pro Glu		
500	505	510
Lys Ala Gly Gln Lys Ala Ser Leu Ser Val Ser Gln Thr Gly Ser Trp		
515	520	525
Arg Arg Gly Met Ser Ala Gln Gly Gly Ala Pro Ser Arg Gln Lys Ala		
530	535	540
Gly Thr Ser Ala Leu Lys Thr Pro Gly Lys Thr Asp Asp Ala Lys Ala		
545	550	555
Ser Glu Lys Gly Lys Ala Pro Leu Lys Gly Ser Ser Leu Gln Arg Ser		
565	570	575
Pro Ser Asp Ala Gly Lys Ser Ser Gly Asp Glu Gly Lys Lys Pro Pro		
580	585	590
Ser Gly Ile Gly Arg Ser Thr Ala Thr Ser Ser Phe Gly Phe Lys Lys		
595	600	605
Pro Ser Gly Val Gly Ser Ser Ala Met Ile Thr Ser Ser Gly Ala Thr		
610	615	620
Ile Thr Ser Gly Ser Ala Thr Leu Gly Lys Ile Pro Lys Ser Ala Ala		
625	630	635
Ile Gly Gly Lys Ser Asn Ala Gly Arg Lys Thr Ser Leu Asp Gly Ser		
645	650	655
Gln Asn Gln Asp Asp Val Val Leu His Val Ser Ser Lys Thr Thr Leu		
660	665	670
Gln Tyr Arg Ser Leu Pro Arg Pro Ser Lys Ser Ser Thr Ser Gly Ile		
675	680	685
Pro Gly Arg Gly Gly His Arg Ser Ser Thr Ser Ser Ile Asp Ser Asn		
690	695	700
Val Ser Ser Lys Ser Ala Gly Ala Thr Thr Ser Lys Leu Arg Glu Pro		
705	710	715

Thr	Lys	Ile	Gly	Ser	Gly	Arg	Ser	Ser	Pro	Val	Thr	Val	Asn	Gln	Thr
		725							730						735
Asp	Lys	Glu	Lys	Glu	Lys	Val	Ala	Val	Ser	Asp	Ser	Glu	Ser	Val	Ser
		740						745							750
Leu	Ser	Gly	Ser	Pro	Lys	Ser	Ser	Pro	Thr	Ser	Ala	Ser	Ala	Cys	Gly
		755					760								765
Ala	Gln	Gly	Leu	Arg	Gln	Pro	Gly	Ser	Lys	Tyr	Pro	Asp	Ile	Ala	Ser
		770					775								780
Pro	Thr	Phe	Arg	Arg	Leu	Phe	Gly	Ala	Lys	Ala	Gly	Gly	Lys	Ser	Ala
		785					790								800
Ser	Ala	Pro	Asn	Thr	Glu	Gly	Val	Lys	Ser	Ser	Ser	Val	Met	Pro	Ser
		805					810								815
Pro	Ser	Thr	Thr	Leu	Ala	Arg	Gln	Gly	Ser	Leu	Glu	Ser	Pro	Ser	Ser
		820					825								830
Gly	Thr	Gly	Ser	Met	Gly	Ser	Ala	Gly	Gly	Leu	Ser	Gly	Ser	Ser	Ser
		835					840								845
Pro	Leu	Phe	Asn	Lys	Pro	Ser	Asp	Leu	Thr	Thr	Asp	Val	Ile	Ser	Leu
		850					855								860
Ser	His	Ser	Leu	Ala	Ser	Ser	Pro	Ala	Ser	Val	His	Ser	Phe	Thr	Ser
		865					870								880
Gly	Gly	Leu	Val	Trp	Ala	Ala	Asn	Met	Ser	Ser	Ser	Ser	Ala	Gly	Ser
		885					890								895
Lys	Asp	Thr	Pro	Ser	Tyr	Gln	Ser	Met	Thr	Ser	Leu	His	Thr	Ser	Ser
		900					905								910
Glu	Ser	Ile	Asp	Leu	Pro	Leu	Ser	His	His	Gly	Ser	Leu	Ser	Gly	Leu
		915					920								925
Thr	Thr	Gly	Thr	His	Glu	Val	Gln	Ser	Leu	Leu	Met	Arg	Thr	Gly	Ser
		930					935								940
Val	Arg	Ser	Thr	Leu	Ser	Glu	Ser	Met	Gln	Leu	Asp	Arg	Asn	Thr	Leu
		945					950								960
Pro	Lys	Lys	Gly	Leu	Arg	Tyr	Thr	Pro	Ser	Ser	Arg	Gln	Ala	Asn	Gln
		965					970								975
Glu	Glu	Gly	Lys	Glu	Trp	Leu	Arg	Ser	His	Ser	Thr	Gly	Gly	Leu	Gln
		980					985								990
Asp	Thr	Gly	Asn	Gln	Ser	Pro	Leu	Val	Ser	Pro	Ser	Ala	Met	Ser	Ser
		995					1000								1005
Ser	Ala	Ala	Gly	Lys	Tyr	His	Phe	Ser	Asn	Leu	Val	Ser	Pro	Thr	Asn
		1010					1015								1020

Leu Ser Gln Phe Asn Leu Pro Gly Pro Ser Met Met Arg Ser Asn Ser  
 1025 1030 1035 1040  
 Ile Pro Ala Gln Asp Ser Ser Phe Asp Leu Tyr Asp Asp Ser Gln Leu  
 1045 1050 1055  
 Cys Gly Ser Ala Thr Ser Leu Glu Glu Arg Pro Arg Ala Ile Ser His  
 1060 1065 1070  
 Ser Gly Ser Phe Arg Asp Ser Met Glu Glu Val His Gly Ser Ser Leu  
 1075 1080 1085  
 Ser Leu Val Ser Ser Thr Ser Leu Tyr Ser Thr Ala Glu Glu Lys  
 1090 1095 1100  
 Ala His Ser Glu Gln Ile His Lys Leu Arg Arg Glu Leu Val Ala Ser  
 1105 1110 1115 1120  
 Gln Glu Lys Val Ala Thr Leu Thr Ser Gln Leu Ser Ala Asn Ala His  
 1125 1130 1135  
 Leu Val Ala Ala Phe Glu Lys Ser Leu Gly Asn Met Thr Gly Arg Leu  
 1140 1145 1150  
 Gln Ser Leu Thr Met Thr Ala Glu Gln Lys Glu Ser Glu Leu Ile Glu  
 1155 1160 1165  
 Leu Arg Glu Thr Ile Glu Met Leu Lys Ala Gln Asn Ser Ala Ala Gln  
 1170 1175 1180  
 Ala Ala Ile Gln Gly Ala Leu Asn Gly Pro Asp His Pro Pro Lys Asp  
 1185 1190 1195 1200  
 Leu Arg Ile Arg Arg Gln His Ser Ser Glu Ser Val Ser Ser Ile Asn  
 1205 1210 1215  
 Ser Ala Thr Ser His Ser Ser Ile Gly Ser Gly Asn Asp Ala Asp Ser  
 1220 1225 1230  
 Lys Lys Lys Lys Lys Asn Trp Leu Arg Ser Ser Phe Lys Gln Ala  
 1235 1240 1245  
 Phe Gly Lys Lys Lys Ser Thr Lys Pro Pro Ser Ser His Ser Asp Ile  
 1250 1255 1260  
 Glu Glu Leu Thr Asp Ser Ser Leu Pro Ala Ser Pro Lys Leu Pro His  
 1265 1270 1275 1280  
 Asn Ala Gly Asp Cys Gly Ser Ala Ser Met Lys Pro Ser Gln Ser Ala  
 1285 1290 1295  
 Ser Ala Ile Cys Glu Cys Thr Glu Ala Glu Ala Glu Ile Ile Leu Gln  
 1300 1305 1310  
 Leu Lys Ser Glu Leu Arg Glu Lys Glu Leu Lys Leu Thr Asp Ile Arg

1315	1320	1325	
Leu Glu Ala Leu Ser Ser Ala His His Leu Asp Gln Ile Arg Glu Ala			
1330	1335	1340	
Met Asn Arg Met Gln Asn Glu Ile Glu Ile Leu Lys Ala Glu Asn Asp			
1345	1350	1355	1360
Arg Leu Lys Ala Glu Thr Gly Asn Thr Ala Lys Pro Thr Arg Pro Pro			
1365	1370	1375	
Ser Glu Ser Ser Ser Ser Thr Ser Ser Ser Arg Gln Ser Leu			
1380	1385	1390	
Gly Leu Ser Leu Asn Asn Leu Asn Ile Thr Glu Ala Val Ser Ser Asp			
1395	1400	1405	
Ile Leu Leu Asp Asp Ala Gly Asp Ala Thr Gly His Lys Asp Gly Arg			
1410	1415	1420	
Ser Val Lys Ile Ile Val Ser Ile Ser Lys Gly Tyr Gly Arg Ala Lys			
1425	1430	1435	1440
Asp Gln Lys Ser Gln Ala Tyr Leu Ile Gly Ser Ile Gly Val Ser Gly			
1445	1450	1455	
Lys Thr Lys Trp Asp Val Leu Asp Gly Val Ile Arg Arg Leu Phe Lys			
1460	1465	1470	
Glu Tyr Val Phe Arg Ile Asp Thr Ser Thr Ser Leu Gly Leu Ser Ser			
1475	1480	1485	
Asp Cys Ile Ala Ser Tyr Cys Ile Gly Asp Leu Ile Arg Ser His Asn			
1490	1495	1500	
Leu Glu Val Pro Glu Leu Leu Pro Cys Gly Tyr Leu Val Gly Asp Asn			
1505	1510	1515	1520
Asn Ile Ile Thr Val Asn Leu Lys Gly Val Glu Glu Asn Ser Leu Asp			
1525	1530	1535	
Ser Phe Val Phe Asp Thr Leu Ile Pro Lys Pro Ile Thr Gln Arg Tyr			
1540	1545	1550	
Phe Asn Leu Leu Met Glu His His Arg Ile Ile Leu Ser Gly Pro Ser			
1555	1560	1565	
Gly Thr Gly Lys Thr Tyr Leu Ala Asn Lys Leu Ala Glu Tyr Val Ile			
1570	1575	1580	
Thr Lys Ser Gly Arg Lys Lys Thr Glu Asp Ala Ile Ala Thr Phe Asn			
1585	1590	1595	1600
Val Asp His Lys Ser Ser Lys Glu Leu Gln Gln Tyr Leu Ala Asn Leu			
1605	1610	1615	

Ala	Glu	Gln	Cys	Ser	Ala	Asp	Asn	Asn	Gly	Val	Glu	Leu	Pro	Val	Val
1620					1625					1630					
Ile	Ile	Leu	Asp	Asn	Leu	His	His	Val	Gly	Ser	Leu	Ser	Asp	Ile	Phe
1635						1640				1645					
Asn	Gly	Phe	Leu	Asn	Cys	Lys	Tyr	Asn	Lys	Cys	Pro	Tyr	Ile	Ile	Gly
1650					1655				1660						
Thr	Met	Asn	Gln	Gly	Val	Ser	Ser	Ser	Pro	Asn	Leu	Glu	Leu	His	His
1665					1670				1675			1680			
Asn	Phe	Arg	Trp	Val	Leu	Cys	Ala	Asn	His	Thr	Glu	Pro	Val	Lys	Gly
1685						1690			1695						
Phe	Leu	Gly	Arg	Tyr	Leu	Arg	Arg	Lys	Leu	Ile	Glu	Ile	Glu	Ile	Glu
1700						1705				1710					
Arg	Asn	Ile	Arg	Asn	Asn	Asp	Leu	Val	Lys	Ile	Ile	Asp	Trp	Ile	Pro
1715						1720				1725					
Lys	Thr	Trp	His	His	Leu	Asn	Ser	Phe	Leu	Glu	Thr	His	Ser	Ser	Ser
1730					1735				1740						
Asp	Val	Thr	Ile	Gly	Pro	Arg	Leu	Phe	Leu	Pro	Cys	Pro	Met	Asp	Val
1745					1750				1755			1760			
Glu	Gly	Ser	Arg	Val	Trp	Phe	Met	Asp	Leu	Trp	Asn	Tyr	Ser	Leu	Val
1765						1770				1775					
Pro	Tyr	Ile	Leu	Glu	Ala	Val	Arg	Glu	Gly	Leu	Gln	Met	Tyr	Gly	Lys
1780						1785				1790					
Arg	Thr	Pro	Trp	Glu	Asp	Pro	Ser	Lys	Trp	Val	Leu	Asp	Thr	Tyr	Pro
1795						1800				1805					
Trp	Ser	Ser	Ala	Thr	Leu	Pro	Gln	Glu	Ser	Pro	Ala	Leu	Leu	Gln	Leu
1810					1815				1820						
Arg	Pro	Glu	Asp	Val	Gly	Tyr	Glu	Ser	Cys	Thr	Ser	Thr	Lys	Glu	Ala
1825					1830				1835			1840			
Thr	Thr	Ser	Lys	His	Ile	Pro	Gln	Thr	Asp	Thr	Glu	Gly	Asp	Pro	Leu
1845						1850				1855					
Met	Asn	Met	Leu	Met	Lys	Leu	Gln	Glu	Ala	Ala	Asn	Tyr	Ser	Ser	Thr
1860						1865				1870					
Gln	Ser	Cys	Asp	Ser	Glu	Ser	Thr	Ser	His	His	Glu	Asp	Ile	Leu	Asp
1875						1880				1885					
Ser	Ser	Leu	Glu	Ser	Thr	Leu									
1890					1895										

<210> 13

<211> 3682  
<212> DNA  
<213> Homo sapiens

<400> 13

gaattcctgg tggagaacag cacatgtaca gatggggtga gaacagcata cgtacaggta	60
ggggtaagct ggtgctatat gagaaagcat ggaataagtt attaagtttgc acctgcttgg	120
gaactgaggg gcaggggtgtga gggatgaagc aggagtaggt aggggctaga tcacaaaaga	180
tctatgccag tgtttctcac agtgtgattc ccagcccagt agcatgatat cacttggat	240
cttgtagaa atacaaattc ttatacatca ccctggacta gaccacctga ataagaaaag	300
ttgggcatga ggcctacaaa tttttaaaaaa agtcatacag gtgattgcaa tgcattgctaa	360
attttgagaa acactcttg ctgtggtttgc aatatttgc tccttccaaa attcatgtag	420
aaaccatctc caatgttata gtattaagag gagggaccct tggagctga tcagatcatg	480
aagtctcattt tcttataaaag gggattaaaaa gccttggccc ttttaccctt tgtccatgt	540
aggacacagt gttggaagca gggactgggt tctcaccaga aacagaacct gccagcctct	600
tggctttggaa cttctcagcc tccacaatttgc tgagaaataa gtttctgttg tttataagtt	660
aaccagtctc aggtattttgc taatggcagc acaaaggggc taagaaactg ttctatgccc	720
taacaagaaa tgtggtcact ttcctgaagg aaatggggat atatataaaag atgttatata	780
agactcgtaa tatttatttgc gaaggcttgc tctgcaagca aggtggaaga gcaacatgaa	840
ggaagcgtgg tggaggtgag aggactggag gttaagttgg tagggagata cagggaaagaa	900
gcttatgaca cttgagttaa aatgttagcat ctttcctatg tgttagggctc ataaaaatgt	960
atagtctaag atagaacaca gaatactcta tgaatcctgc ccacaagggtg ttggtaatct	1020
agattcactt ttttttctg ataatgccat ccatatgtat ggagcgtctc ctactgtatg	1080
ccagagtgac tctggaatcg gtttgggtga tctagacaag accataagga gagtccccctt	1140
actacctctt ctccaggggaa gggattcaag ttgaactagt acttcagaga ctgttttagta	1200
atatcatgca tgaaaggtga tggtaggac agaaaaataa atggattgca tcataattcc	1260
tcaggttctc caaatatgtg gtggctcaa accatgtgaa ttggcttgca catcctgttt	1320
gggttgcgtg tcagcagttg agatctgagc cttatttgc acagtgaaac agtgagagac	1380
ctgcccttca agagctgttt ttcagctagg aatagaaaag ggccaggctt gactccttctt	1440
tctgctggat cttgcttctt ctcagcaata gaagtagacc tgccttcata gctgtagaga	1500

aaagggtgccg	gtaggcgggc	aggtgagcct	gtggataatc	ctggagtaaa	ggttcaatag	1560
accttcaagt	ctatcctaca	ggattcggag	tgaggggaga	aaaaaggaga	cgttctctg	1620
gctgagagag	gaagagaaaa	aaaaatccc	gatatctgac	agctatatct	tcccatcacc	1680
accttcctct	aaacccatgc	ctctctgttt	agtaggacat	aaaatgaaga	gtgacccacc	1740
ccccacccccc	agcccatccc	ccgtttgtag	gtgtgcttc	aatgaaaata	agtcgggttt	1800
catggacgga	aactagagca	gtgaaaata	gatgcaagac	ttgttgagca	tacaaatcat	1860
ttccccctta	gtctccaagg	gaggaaaaaa	aatccctctt	actctccttg	cagcctgtgt	1920
tctgcattct	ggagaggaag	ctgaggctgg	tcctcaggcg	ctcctccgc	cgttcccgca	1980
gaaaaactttt	ctcgcagggc	ccgctccgtc	catcccgccg	ggttccaaga	cgggtgggcct	2040
cccggtggct	ccttccttgg	gcaagggccc	agaccccgcg	acgcgcctgt	ctctttaaat	2100
tccagctgcg	cggctggaa	acagcgccac	tcgcccggca	ggccggctgg	aggctgaaga	2160
cgagagctgc	gctttcgctc	ccggctgccc	gccgcggaga	gctgggctcg	gcccgcgggc	2220
tgctaggtgg	cggcggcgcg	gggcggggag	gcgcggcccg	gcggaggagg	gaagaaagag	2280
cgagccgggc	cgggagagggc	gccgcgcccc	gtcccgcgccc	cggtcccgca	cccgcctctca	2340
cgccgcggcaag	cagtttcttt	ctgggtgaca	agaatgtgcc	tcgggtggtt	tttctttttt	2400
ttctccatct	ccttaagacg	atttccatag	taacotgatc	aagtggctca	aaatcgcaaa	2460
cctgaggatt	tccgcggccc	gccggcaaga	cctcggccag	gtaacgctgc	gatctcctcc	2520
tcttccattt	caaaccgctg	cgctccttgc	aaagttcctt	ttgtggaaaa	tcgcccagcc	2580
caagggagcc	cgggttattt	gcaacagcgt	gttcatttcc	aggtgcctgt	cacgggtctc	2640
cctccctgctg	cttctccagg	acccatgatg	agattatttt	taaaaaattgt	ttttggtcgt	2700
ctccccccgccc	ccctccctt	ctttatTTT	ttcctcttgc	ctgcacttt	ctcggtttt	2760
ccctgacac	tactgatggg	ggtgcggggg	gacgtcgggg	atgggggtgg	ccagcgcggt	2820
cctggagtg	gcgggttcgg	atgggctggc	tgcggtggc	cactttggc	atctcggcgt	2880
ggcctgcgcc	ggggcacgg	ggagggctgt	cagcgccagg	gcggcgaaac	ccgaggtctc	2940
cagacgagtg	agggagggat	gcaggcttgg	gggtgatgga	gctgtgggt	ggtggctgg	3000
gagcgtccat	acatcatagc	tctccttccc	actccccgc	ccctttcgg	gattctctct	3060
ttctctttcc	ccgtcctcat	ttcttttttc	ctttactcac	cactcgcttc	attctttcc	3120
ttccatttcc	tcttttttcc	tcccctcatt	tccttttttt	cctttccctt	ttaaagaaag	3180
ggaaatcggt	tgtaaccctt	tcgttctacc	aacgtggaat	agctgtgaaa	cctgcagcgt	3240

ggtcacacctca gcctggtcgt tttcagaccc gtcctcatcc atcaacatat ttgtttcccc	3300
agtctattga tctccctgaa ttctacagaa atgcattcta agctaggcgc ctgtatgtca	3360
gaatcagttc tgccaggtagc ttccgtgctc caagtatgac atgtattgta agggctgcat	3420
ctgtttaaa cccacataag ccatgggtat aaataaatgt agcttgaaa aaaaatctgg	3480
ccttattcta gataaaacttc cctcttaaat tactgatata ctcttctccc tctttgacat	3540
ttaatttttag gaaagttggg agacaggttc ttgtcctcca gtttttaagg agcaggcaac	3600
ttctatttatac ttaattttct cgtcttgaa catcaactcac gtttgcacta cccagtcagt	3660
ggaaccgagtg ggtcataatt aa	3682

```
<210> 14  
<211> 3904  
<212> DNA  
<213> Homo sapiens
```

<400> 14  
cctgcattat tgtttttatac tgacttccaa ttttggtgtt ccctgggtgg gtggggtttc 60  
ctgacacatt tacaagatgc ttttggcagg ttggctggaa tttgaaggca catttaattg 120  
taggtgcaat aaaatattca ttttctcttg ttcttggttt gagatgtcat gccctttgg 180  
tcacttatat tttgggtgtga ctgtgtgtgt gtgtgtatgt gtttgggtga aggatttAAC 240  
aaagtctgtt ctaactgtca tgtgatttga agttaaaagg tatgttagtg acaagccaca 300  
aatttctctt atttatacgta cattgatcct gaaaccattt tttcccttgtt gatttcttct 360  
gtgcatggat catttaacga aagggtggca atgatgagct atttttttat aataggaaaa 420  
aaattcctca agtttactta ccaagtcata tttttataca gagggatttag caaatattc 480  
tgatctaata tttaataga ctgaattgct gaccactgct aattaccaag aatatatTTT 540  
cttaattctg aaattgctgt acctctcaag ttgtctggag gactccaagt gacccaactt 600  
gtaactcatg gcaacaggaa gtgggtgttc tgggtgcaag ctgaagtgtg cacatggacc 660  
cgtactttgt tagcactcgg ggacttgata tggaaagaat taatgtactg gttttttgt 720  
atagatgaat gttaactttc tgacattagt cagaactaca tctcccaagc cttgtttgc 780  
agtgtctgtc ccttgctct tcacttacag taagtcctta cttaactgac ttgataggtt 840  
cttggaaaact gcaactttaa gcaaaaggaa gtataatgaa acacttttat cacaggctaa 900  
ttqgtqaaaa caagacttaa gttcccatgg catattctg gtcacaaaaa catttccaaa 960

cttctcaaaa cacttcaata ttaagcattc aaatacatgt aaactatgt tatatgtaa	1020
aaaggttact ataaaccaga tcaatattta cccaattatt taagttcagg gtcttagtg	1080
gctggagcct atccgagtag ctcagggcac aaggcggaa ccagccctag acaggacacc	1140
atcctgttgc agggcacgtt cacacatgcc cacacgcagg ctgggaccat ttacatgtgc	1200
caattcacct accatgcaca tctttgagac gtggcaggaa gcaagagtac ctggagaaaa	1260
tccatacaga tatgggaga atgtacaaac tccaccaga cagtggaccc agccaggaat	1320
caacatttgg gcaacattat aatgaaacga agttgaatga aatgatgtcg ttccacgacc	1380
tgtgtactt gaggggtgtt ataaaattct cagaagacag aggtttaatg ctattttt	1440
aatagaaaaat aacttataga gaagtgtgca catgtgactt tgtgtgtgc aggaatcatt	1500
agcatgagaa tcagacgtaa gaggtggtgc caacatgagg aatgttggaa ttccaggagc	1560
tgtggatgga agtagaagcc agaaggccag ggttaggttc ctacttctta ctgtttcagt	1620
tattgcagtg ttggcctgtt tattcacaga tgtcacctag ctttgtttc tcaagaagaa	1680
aaatgagcat aatcttcct gttatgaatt cttaaacaca caggacataa ccacagacac	1740
agaggtgcac atatgttagca gtaatggata cttaatgata cactcgagg aaacagaaaa	1800
gacttctgaa tagagactgg agataacttcc ttggaccatt gatgaatggg caatgatgca	1860
tttttgtctt ccattcagaa ggctaataata ttgctctcta tgttctatgg ataaaggcag	1920
tatatgctca aggatgaatc acataatatg cataataaat ccagcaagca ttacccttt	1980
acttatgtga ctgcaagtag gaatacattt cccccactct aaccatgtaa gatttcttc	2040
ccttctccca ttttgtaaagc aaaagtaagt tcctgaaagg ttaaatggac ctcaggatgg	2100
aaaaaatccc cagagctatc tttctgcaca gacttcattt tttctcccaa gtctgactgt	2160
caactgcgat atctgatatg aggctctggt gctgatgttt ccataggtca tcatccttcg	2220
gtgtcccaga tgaagtctca ggtcgAACat tgcaatagca cagattctga atttaatgca	2280
tcattaaagt tggttatgta acccaatggc cttgttaaac tccagatttt taaaattata	2340
tgtatTTact attctcttat tttagaatga tctcacaatg ttcacaagaa ataagcccag	2400
tccctgcaaa gactttaaaa gctgcttggt cacatcatta gattgtacaa cgcttgcata	2460
atgacacttt ttgctaatct atgcaacattt tttgtaaacaa ttgtgcacat tttaactact	2520
tcagataatc aggacctaga gacttcaaga tctgaaagca ttgctggtga catagagcaa	2580
aaactttctt gagaatagga agtcagtgtt ttgacaagtg atttataaca gttcaggtat	2640

agccaggaag	gtttcaaaca	aacctaagt	attatttctt	tcatcttgat	tagtatata	2700
ttatatgtga	tctatttatg	tatattaata	gatTTTggg	tcttatagcc	agcttcatt	2760
tttctctatt	ggaaaagatc	taagtcccc	tccttccttg	gtggctttg	gtaggttgt	2820
agacaaaaca	ttgaagaatc	aatggtacct	tttatacatt	aatactgcc	atatgaccat	2880
aaaatcatat	tttttggaa	tttattcccc	cgtcaaaag	aagcatttgt	tattgaacac	2940
agtcttatgc	taccttatta	agatgtatca	aacaccctga	ttgatcaaaa	acaccctagt	3000
ccatTTtaag	gcagtattgc	ccagcaatta	aagatgtac	ttctggagga	gtctttctga	3060
gtttgaattc	agtactcttc	cacgtactat	ataggtgatc	ttgggtaaac	ttcttgagtc	3120
tcagtagccc	catctgtaaa	attgtttag	agaagaattt	ttgtgatgat	tagtgagag	3180
aatatattaa	tgtaatattt	aggagagcaa	ccagcatgt	gcataatattc	attacatatc	3240
aatttctata	ttattgatgt	tcatactgct	gatgttgaaa	tgcacaggaa	ggccacagtt	3300
atTTTctgtt	tagattgatt	tttcttttaa	agtctgaaca	taaactgtaa	tactgtgctt	3360
atTTatgtag	gaactgtgat	ctcgtctcct	cctttccca	tctccccctc	tctaccttag	3420
ttttcctta	tagtctcaag	ctgaaaacaa	tgaccaggt	cctaagagat	aagaatactc	3480
tttctttga	actcatggca	ttagcagtga	cctggatgag	attggaggct	attattctaa	3540
gtgaaatagc	tcaggaatgg	aaaaccaagc	attgtatgtt	cttacttata	agtggagct	3600
aagctatgag	gatacaaagg	cataagaatg	acacaacaga	ctttggagac	ttggggaaag	3660
ggtgggaagg	gggtgaggga	taaaagacta	caaatagggt	gcagtgtata	ctgcttgggt	3720
ggtgggtgca	ccaaaatctc	acaaatcacc	accaagaac	ttactcatgt	aaccaaacac	3780
cacctgttcc	ccagtaacct	atggatataa	aaaaattaaa	aaaaagaaaa	aaagaaaaact	3840
ctttttgca	ggggcgaggt	aaaggtaag	agggcatccc	atTTTgagt	ttctagaaaa	3900
gctt						3904

<210> 15  
 <211> 2119  
 <212> DNA  
 <213> Homo sapiens

<400> 15

ctgcaggaag	cagcagcaag	gtccaggag	cctctaattt	aaataggaga	agtcagagct	60
ttaacagcat	tgacaaaaac	aagcctccaa	attatgcaaa	tggaaacgaa	aaaggtaagt	120

gtttgttaca tcattatgac acaagtccaa catgagtctt gtgaattgca tgctaaatct	180
aatatggag cagcgtaaca actttgggcc tagagatgtt atcagtggag tttctttatg	240
tttcctaact gtcccctcct gactgccagc tttcttatct gaagaacatt taaaacaaat	300
aaactcattc attttaaagt agttagttat atatgcaagt acaaatactg tttctcaaaa	360
acaggtcctt ccaaattgcat gtaaatcaca ttttcttatg tcttttatg ttttgaaaa	420
tgtatcctga aatcataaaag ccatattgaa tttatctgaa tccttaactt cagttaggt	480
aagagccata agtgttttg acaattaagg ttggagcatc aaaatttcaa acataattac	540
agtaggtttt tatcttgca agcagcagat cccagagata ttatgaccc agtttcccc	600
aaaagacaaa ttattcatat ttgtttgtt ttcttgaatt agtgcataat ataaatatca	660
aatcacaaaa tcaaggacat taaatgaaag tgtctgtta aggcatatta taaatgaatc	720
ataagccaca cagttctctg tgatgtacga agtggcatt taaagaggtg ctgatttgat	780
gcttgtcact gagtagcaga gaggacgggg atgagtatgt gtagttaca cctcaatcat	840
gaggaagtga agaacttgc ctgttataag tagtatggct gtgtgaggaa ctgggtgtt	900
ctgctggatt ttgaggaagt attttcaaata caatagaact tcaaactttt cttcagagtg	960
ttgggctcta catggaaaaa cacatgaaat taaaaagtgg cacaatgtt tagtttagtag	1020
aacatctggc taattggat caaataattc aaccatgtgg gaacgtttt gctcaaaata	1080
gataattgtg aattgttca tataggcaaa tgattagaca acttcctt cctcaaatgt	1140
gaacggacag atgtgatcta gaagcaagac actctttgt gtaaatattc ctttggcct	1200
aaagcaaaag tggacagact ttaaacaccc gagagcagag cagtgtgtt taagattgca	1260
atatcttaag ctcttgagtt aaatggaaaa tgaaaaacaa aagtgtatata ttggaagtt	1320
ggaatgtttt cttaaaaata taaaataaaa ttttagattt aagatcaca gaaatattac	1380
tgaagactta tactcttcct ggggctaagg gaggtgacag tcgctcatca gaaaaaaaaa	1440
aatgcctca ttccctaact tttctaaaaa atataataca agttcaggt aataacttcct	1500
gtatatgtgg gaaatttcta ggggaagcta acaggcttag aaataaagat gtgttaata	1560
gactacaaa gtgtccaatt aagcaacacg ataccaccgt tattgatatt ctagcaagaa	1620
attactagca atgtttgtaa atagacttag aaatgcattt gatgaattaa cactttata	1680
tcttaattta tctgaatttt tctgtatgt gaaaatgttt tatttaactt atttctggca	1740
tctatttagta aaattctgat gatatacaag cattaatatt tttccatggc cactcaattc	1800
atacataacct tccctatcta tgcttagaaag gcagtgc当地 attagatgt agcaatattg	1860

attataacca caaggtggag acagatgtca tgtaatatgc agtctgctca tataaagcac	1920
attttcttag acaagagttt tcatacgata taataaagac atctggaatt tgtcttgat	1980
gcaatatgaa atttgctatt aaacgtggag ttaaaaactt atgtcaatag atccaataac	2040
aatgttcata aattaatcat tatgtcatgc tgtatttcca aaatactatc taaaattata	2100
agagcaaacg aggtaataa	2119

<210> 16  
<211> 2103  
<212> DNA  
<213> Homo sapiens

<400> 16	
gtacatttt taataaaagat gtttgtttta acttttgaa tatgaagatt tctagttcta	60
gaataatgtt tataaaaata tacaaatcca tctggtgatg agttgaccc tcatacact	120
agttgcata tataacttgg gtgtgaccaa gcaaggtag agttaagaac ttttaaaact	180
tactgtatta tattgataga actcagaaag tactaacttg aatattatta ttcttaattgc	240
ttttccctt tagttattaa aaataagaat acttaaatttata ataacaagat ctttactgg	300
caggattaac caaatttatct gtaatgtgtt cctcgatgc ttttaagtgg aaatataact	360
tatacattct ttaacaactc tgagaggatg agttacataa atcagttcag gaatctatag	420
aatctgtaat acatagtaaa ggtttattca caattaaaac aatttcactt ctatattaaa	480
aaaacaaatt gttgaaagta cagtggctt tcataatgtat gatttgtaaa acaaattagc	540
tttttaaag tcatgtgacg cttaatgaga agaaatcagt agagaattac aaactgcact	600
tcaaaagata catctaataat cattttataa atgaaatttg aaaaaatagt gtgctcggtt	660
tacagtctca taaaatgaat taaaatatca gcacacattt tagtaggtt tcattggcag	720
agaaggctga aatagaaaacg ttacaatggg atgcactgcc atctgaacat tatgtcgaag	780
tggAACGCGG aaacatattt ctcagaacaa gtggtaaaat gaaaacagca tcatttgtaa	840
agcatttctt ttgagagtgc ttcatgttct tctcctgatg acctgccatt cagaaactgaa	900
caatgaataa tacactctga caccacattt tgtcaatttg cccagaacca tatgagagta	960
ctctagacag atatatgttc cgaagtaaac cgaatacctg ttaactgtaa atcaaatctt	1020
gtagaaaacca tgccatggtt cctttggaca tatactttgc atgcctgaag caagttacct	1080
taagaaaatca ttctttgtt ttacaaaact tgtattaaaa aattaaaaat gcaaaaaaagc	1140

ttaatattat taggaattta tccatagctt tatttggaaat ccagtttctt tattatgatc	1200
tataaacatg catcatttga tggagttcct tagtggagag gtgttttcc atgttgctaa	1260
gaaacatgcc ccagcaccag aagggatact acctaccatc ttttgcctt ttctcaccgt	1320
gattcttaca ttgtacctgt ttactcaactg aacagggctt ctttcctttt gtcttagattc	1380
taatcaggtg tcttctggtg tggaaagcttt ggcttttatt tacacacaac acagaattaa	1440
taagatagat gccaaggatt tagcaacatt ttaattcaac attatacagg tatcagagtt	1500
aatgagaatt atgcatttagt cttaaattt gggcagctt ttcagctaaa acatagatgt	1560
ctagctctta aacactttgt ttttttaatt actctgaaat tacaataaaag tcaaagaact	1620
gaactgtttt ctttcaagc cagtgc当地 gtgcttttagt tattattta ctggatct	1680
aattatgcat ttaatgctt tattacttaa tacttatata agcctaaaat acgttgtttaa	1740
tgtcataatt tcagggattt tagtattctt tccatgagtt accataacta ggtgc当地	1800
tgtaaatata cgtatatac tatatctata tatttatata tatgtatata tcaatttata	1860
agactaaata gacttggcca tatgtgttgc tggttatgc atacatgc当地 aaatattgag	1920
gtgtccacaa agtataatag cctgtacata aattacatac tggctggtaa gtgaatgtaa	1980
gcttctctaa attgtacaac tctccacaga gtggcactct aatattgcaa aggtacaata	2040
taagcatgtg cagaatgaac agctttcta ggatccctat aaaactccac cccatgtttc	2100
tgt	2103

<210> 17  
 <211> 4260  
 <212> DNA  
 <213> Homo sapiens

<400> 17

aagtttc当地 ccagaggggc acttgc当地 agtctc当地 tatgaggagt	60
ctatcaacac ctgctggag gtgtctc当地 gtcaggaggc acgggggtca gggaccact	120
tgaggaggct gtctgtccct tagcgaggct agaacactgt gctcgagat ccgctgctct	180
cttcagagct ggcaggcaag agtgttttag tctgctgagc ctgc当地 agccgccc当地	240
tccccccaggct gctctgtccc agggagatga gagttt当地 tgtaagcccc tgactggggc	300
tgctacctt ctttcagata tgccccgccc agagaggagg aatctagaga ggcagtc当地	360
ctacagc当地 tttgccaagc tgc当地 gggc tctgccc当地 caaaattcc cagc当地 ggggttt	420

gtttacattg tgagggaaa agcacctact caaggcctcag ttatggcagt tgcccctccc	480
ccccccaagc tccagggtcc caggtgtcct tcagactgct gtgctggcaa tgagaatttc	540
aagccagtgg atcttagctt gctgggtcc acaggggtgg gatccactga gctagaccac	600
ttagctccct ggcttcagcc ccctttccag gtgagtgat ggttctgtct cactggcatt	660
ccaggtgcta ctgggtatg aaaaaaaaaa ctcctgcagc tagcttggtg tctgcccagt	720
tttgtgttg aaactcaggc cttgggtggt gtggacaccc aatggaatct cctgggtgc	780
atgttgaa gactgtggaa aaagcatagt atctggctg gatacgctcg tccttcaagg	840
cacagtcct catgacttcc cttggctagg ggagggagtt ccccaacccct ttgcacttcc	900
caggtgaggc aacacccac cctgcttctg ctcaccctct gtggctgca cccactgtct	960
aatcagtcac tgtgagatga gcctggtaacc tcagttggaa atgcagaaat cacctgcctt	1020
ctgtgttgat ctcactggaa gcagcagact ggagctgttc ctattcagcc atctttctca	1080
ggtcataatc atagattttt aattgatccc agcaacatgg attagtaaac agcatatttc	1140
caagtgattt ttttttattt taaggtcaaa tctacaaaat attatagtgt tatcaccact	1200
taaaattatt actggtgata ctatgtttgt ctctattcac attttattgc tagaaagaat	1260
tataatttgt agataataat agttatttga aatgtattac atatccttt acttttaaga	1320
agaggtgact taattatcta ggtatacaat tatttgagg atactaaatg tcatgaatag	1380
caaatttatac atattgcttt cctaggtgaa gaccctgaaa caagaagaat gagaacagtt	1440
aaaaacatag cagacttgag gcagaattta gaagagacta tgtccagtct tcgtggact	1500
cagataagcc acaggtttt ttcaattttt catatatttgc agccaataaa gaaaaaataa	1560
ttacaaacaa acatthaact tttcttataa tgacagagat gggatttcag tttccctta	1620
ctatttctc cttgtttta tatcaaatttgc attggtaatt atccttaaac tgagaattca	1680
cagtatatac ctatttatct tttatctcta tctctatctg ctatttatgt cttttcagt	1740
ataatttcca gtactgcaac taccaccatc actgttaagt ggatttgcata tacctgtcct	1800
agaaaacagt ggcacaagtt gcacttgaaa tgcacatctggg cagggtagta gggagacatt	1860
caaacataat tgttagttaac tttcagaata ggtctggaa ggttacagt gatggatggat	1920
ttgttgaaaa tgtaaaacaa tatgttgattt tacccttgcgtt gttactgtgg cttttttttt	1980
gaaaacaaac gaaaagctat aaaatgtatg cccctttcca caatttgcacc tcaaaatgaa	2040
tatagagttt agcttcggg aagatgacgt gtttataaga gatgaccctc aactccagcc	2100

ttttctgtct tcatgcattc tagattatgg ccctaagtga accagagtgat agtattttct	2160
ccattttatt tgacagcacc ctggagacaa catttgacag cactgtgaca acagaagtttta	2220
atggaaggac catacccaac ttgacaagtc gacccacccc catgacctgg aggttgggcc	2280
aggcatgtcc gcgacttcag gcgggagatg ctccctccct gggtgctggc tatkctcgca	2340
gtggtaccag tcgattcatac cacacagacc cctcgagggtt catgtataacc acgcctctcc	2400
gtcgagctgc tgtctctagg ctgggaaaca tgtcacagat tgacatgagt gagaaagcaa	2460
gcagtgacct ggacatgtct tctgaggctcg atgtgggtgg atatatgagt gatggtgata	2520
tccttggaa aagtctcagg actgtatgaca tcaacagtgg gtaagtaacc ctgttctccg	2580
tcagcattgt gtgaagaggg gaggtggctc actataatgc attcactata aacaaatgtg	2640
taagtttgcc cagaaagtca tgagaacata tgagatatct gaggttattc agagtgttga	2700
aggcccttc ctctgctcat tcatggagag taaagaatcc aagatttcta taaattcatt	2760
ataagccgct aagttttct gttgttgaga gaaacacatg tggcttctgt tttcagagt	2820
gattttcaca tgcttcttaa gtaacagatt ttgttagttaa ggacgtggga aggagacagg	2880
aggagtttg ctgatttgct tgatttttt tttctttttt agcttggtag aagcggcctg	2940
taactgctt gagaaacaaa tattttctta ctgtcttcaa ttatgcattcc ccaatttaac	3000
ttgagggaaa aatcactttg gagttgaaag tttcactcta ttcattttct tttgatggta	3060
tcagatttca atacatctca gaccctgttt ttcttctgtg tccttattaca ttccaaaaca	3120
tgttgtgatt gtaaaactct tagagtataat taacaatttg ggatatttgg cataatcaga	3180
gaataggtcc aaaaggagggc aataggatata tctattaata attgtatgg ccatttttag	3240
catttcctgt tatgtactat gctcttgcata agtgctttga agatagtgtt ttactttcc	3300
ttccccaccac cagcaatgtt tatgaggttag atgtttttat acatgttcta tggataagga	3360
aactgagtct aattggcccc ggctggaaac taacgctagg gaaacggcag acctgcattta	3420
gaactcagct atgtctgact tcaaacacag gctcagtaat atgtggaaaa gctttccaat	3480
taactttgtc tataaacttt gtgtgagtct ggattttgac ttactcttgc tctttacgca	3540
tctgagagga cccatgtagg aaataattct tctatataag tgacccttcc tgacttcatt	3600
catgaaaagc ttatgtttga agggtgacac gacctaaaaaa agagtacaaa atagcttttg	3660
attacattta tagcttgct ctgatatcct aatacctact agtccattcc tggatccac	3720
cctacctgac tttctaaaaaa tttagaatta tagagactaa ttatgattaa ttaagatagg	3780
ttgttgtca gttgccactg gattcagat gcctagtttgaatctctccc attcactata	3840

tgtggacccc ttcggAACCT aacgtatcca aattagttt tgtcatctag aataaggata	3900
aaattgtacc atttcatga agttgttagg atcatccaca aatttttagtt tgCGCAATgc	3960
ttggcatgtat acaagcactc aataaattta tcattttcct ctttatcatc actattacat	4020
ttattatcat taataaccat accaattttt ggttgggtt agttataatt atcatttttg	4080
tatgtattta acatagccta ggaggcaatg cccagttcag aaaacataat ggcaaAGCAA	4140
gagtgtctaa ggcacactct ttctccatc tctctttct ttcttctcca ttctttcac	4200
tctatccccct cttctttttt ttttctcaat ctcccttagat gtggacatat gtgtgaattc	4260
<210> 18	
<211> 2429	
<212> DNA	
<213> Homo sapiens	
<400> 18	
tgtgggtgtg ggtgtgaAGC atgtgtatgt gtgtgtgtga agcatctccc cacctgtaat	60
gtaagtccat gagtgccagaa ttttgacat attctttacg tggtgagttt taacaaatgt	120
ttgtggagtg aatgaacaaa ttaatgaata taggctattt attaattagg caatatagtc	180
acataggctg gcaatcgcat ctaattaaat agagtggtaa atgagttcca gaaagaacta	240
aggtaactaca aggatgttat gaaagagaaa aatgagttat gtgaaaaata ggagacagt	300
ataagagggaa aagaatccc aagtgtggc cacatttga aactaatgac ctattattct	360
attattgtta gctgaaagta gaaaacgtca tgggagggaa tatctgctag tttttggtaa	420
aggatgttgt gatggcagaa ccaagaaatg aacacaagggt gactttggtt tggggacagt	480
gggataatca actctccttg ctccatcagg gccccagact gggctctggc agaggaactc	540
agaacaacgt aaagacctag ataggtatct aataaattgg gacctgtgaa aacagtgcct	600
cttaaaagtgt ggtacctgga ccagcagcag cagcagcagc agccattgaa acttcataga	660
aagacagatt ctcagcttca tccaagactt actgaatttag aatatctcaa ggtaaggcct	720
ggtaatctga gctttaacta gccctcaagg tgattctaa gttcaaggcat cactatatta	780
agttgaacaa atagatgcc a gcctataaa tacatgtAAC gcctAGCATA aatatttcaa	840
cattaaaaat gacatttcat agttcttatt tacccattt gctgtgttct gtcaagataa	900
tgagaatatt gatatgttag aatacactga tgcactaatt tttaaatttag atcaaataat	960
gacttggat acctgaaata aattgggtca gcttggtaga tgcagttttt gagaattata	1020

taagtcattt	ttaaaaagaat	aatttttaact	tgagctgctt	gcataaaatta	aattgcaaaa	1080
aggcatagt	ataaaatccctc	ctattagcag	agatagaagg	tttttaaaaaa	aattacagat	1140
aagtctgaag	gtcttttaaa	atcttatatt	caggaagtga	ctcgggatgt	atatcatttt	1200
aaaatacatg	gtcttaaatg	ttgttagttgt	atgactctt	cagttaattt	aaaatacttc	1260
cttctatgaa	aaattgttcc	aaaaattttt	ctaaattctg	ttatccattt	caagtaggat	1320
aggcaagaac	agatataaga	tactactttt	ttgttcatgt	ttactaaaaaa	aaaaaattact	1380
gtaattgaga	tcatgtaaaa	acatgtttcc	tgtctatttg	tcttaacctt	ttaatccctgg	1440
caccttaaat	ttgacatagt	aggaattaga	agacaattgc	agaaaatgtc	aactggggaa	1500
attttattct	actaaaaact	atgtccatac	aacatagcaa	atcacatttt	aaaggccaaa	1560
aagtctttca	tagcaatttt	tcaagatttt	ttcaaagcat	atcttctctc	tgctccgtca	1620
gcatgccgtt	gattttctg	ttatgcagtc	acataagtaa	ttacatgttt	acatgtctat	1680
ttcactcata	gaacacgaaa	cagttaaatg	tagaataata	tccaatccat	cttttatca	1740
ccagtagcta	gcataactgta	ggaactcaat	aaatatatca	gataaattgt	ggaataacc	1800
atatcagctt	ataacatata	gaaatgtgag	tttaaaaaga	aaacaattat	acatatgaaa	1860
aaatttttat	accattttt	taaagacctt	tcaagatgtca	tacagtttgg	actttccag	1920
tgtttcttgt	atcatgagac	aatagtagac	attgtaaatc	aaaaatagtt	ttctggggtt	1980
gtgtacattt	aaaaaaactg	aatatcatat	ctgttcttag	agagtaatga	tggatattaa	2040
catatcaaag	gtacagagaa	gtcttaaagt	tcaaagtaac	atctgcttaa	ttgtattttaa	2100
ttcagtgctc	catgagctt	tttatcactg	attccctccc	tttttctct	tatgataata	2160
attaacttgt	tcctgttagca	tttaagaaa	tgttgattta	gttgaatgcc	ttcacttctc	2220
caatataata	gcagaaaactc	agaaatattt	atttacccag	aatcatgcag	ctaataatgtac	2280
aaggattcag	gtctttact	tcctattttg	tggttcccaa	ctactttgc	caaagggtctt	2340
ttaaataata	tgaaacatata	tagtgattga	ttcattatag	taaatgggta	aatgataagg	2400
cttgcataaa	ttcactgaca	agaaagctt				2429

<210> 19  
<211> 2926  
<212> DNA  
<213> mouse

<220>  
<221> CDS

<222> (2) . . . (2926)

<400> 19

a agc cac agc acc ctg gag aca acc ttt gat acg act gtg aca act Ser His Ser Thr Leu Glu Thr Thr Phe Asp Thr Thr Val Thr Thr	46
gaa gtg aat gga agg gcc atc ccc aac ctg aca agc cga cct tcc ccc Glu Val Asn Gly Arg Ala Ile Pro Asn Leu Thr Ser Arg Pro Ser Pro	94
atg acc tgg aga ctg ggt caa gcg tgc cct cgt cta cag gct gga gat Met Thr Trp Arg Leu Gly Gln Ala Cys Pro Arg Leu Gln Ala Gly Asp	142
gcc ccc tcc atg ggc gct gga tat tct cga agc ggt acc agc cga ttc Ala Pro Ser Met Gly Ala Gly Tyr Ser Arg Ser Gly Thr Ser Arg Phe	190
atc cac acg gat ccc tcc agg ttt atg tat acc acg cct ctc cgc cga Ile His Thr Asp Pro Ser Arg Phe Met Tyr Thr Pro Leu Arg Arg	238
gct gct gtc tcg cgt ctg gga aac atg tca caa ata gat atg agc gag Ala Ala Val Ser Arg Leu Gly Asn Met Ser Gln Ile Asp Met Ser Glu	286
aaa gca agc agt gac ctg gat gtg tct tct gaa gtg gat gtt ggt gga Lys Ala Ser Ser Asp Leu Asp Val Ser Ser Glu Val Asp Val Gly Gly	334
tac atg agc gat ggt gat atc ctt ggg aag agt ctg aga gcg gat gat Tyr Met Ser Asp Gly Asp Ile Leu Gly Lys Ser Leu Arg Ala Asp Asp	382
atc aac agt ggg tac atg aca gat ggt ggg ctc aac cta tat acc aga Ile Asn Ser Gly Tyr Met Thr Asp Gly Gly Leu Asn Leu Tyr Thr Arg	430
agt ctt aac cga gtc ccg gac aca gca act tcc aga gat gtc ata cag Ser Leu Asn Arg Val Pro Asp Thr Ala Thr Ser Arg Asp Val Ile Gln	478
aga ggc gtt cac gat gtg aca gtg gac gca gac agc tgg gat gac agc Arg Gly Val His Asp Val Thr Val Asp Ala Asp Ser Trp Asp Asp Ser	526
agt tct gtg agc agt ggc ctc agt gac aca ctt gat aac att agc aca Ser Ser Val Ser Ser Gly Leu Ser Asp Thr Leu Asp Asn Ile Ser Thr	574
gat gac ctc aac acc acg tcc tcc atc agt tct tac tcc aac atc act Asp Asp Leu Asn Thr Thr Ser Ser Ile Ser Ser Tyr Ser Asn Ile Thr	622

gtc ccc tcc agg aag aac act cag ctg aaa aca gat gcg gag aaa cgt	670
Val Pro Ser Arg Lys Asn Thr Gln Leu Thr Asp Ala Glu Lys Arg	
tcg aca aca gat gag acc tgg gat agt cct gag gag ctg aag aaa gcc	718
Ser Thr Thr Asp Glu Thr Trp Asp Ser Pro Glu Glu Leu Lys Ala	
gag gga gat tgt gac agc cat ggt gac gga gcc gcc aag tgg aag ggt	766
Glu Gly Asp Cys Asp Ser His Gly Asp Gly Ala Ala Lys Trp Lys Gly	
gct act tct gga ctt gct gaa gac tcg gag aag aca ggg cag aaa gcc	814
Ala Thr Ser Gly Leu Ala Glu Asp Ser Glu Lys Thr Gly Gln Lys Ala	
agc ctg tct gtg tct cag aca ggc tcc tgg agg aga ggc atg tct gcc	862
Ser Leu Ser Val Ser Gln Thr Gly Ser Trp Arg Arg Gly Met Ser Ala	
cag gga gga act cca gct aca gct agg cag aaa acc agc aca agt gca	910
Gln Gly Gly Thr Pro Ala Thr Ala Arg Gln Lys Thr Ser Thr Ser Ala	
ctc aag acc cct ggg aag aca gat gat gcc aaa gct tcc gag aaa ggg	958
Leu Lys Thr Pro Gly Lys Thr Asp Asp Ala Lys Ala Ser Glu Lys Gly	
aaa act cct ctc aaa gga tca tcc ttg caa agg tct cct tca gat gca	1006
Lys Thr Pro Leu Lys Gly Ser Ser Leu Gln Arg Ser Pro Ser Asp Ala	
ggg aaa agc agc ggg gat gaa ggg aaa aag cca ccg tca ggc att gga	1054
Gly Lys Ser Ser Gly Asp Glu Gly Lys Lys Pro Pro Ser Gly Ile Gly	
aga tcg aca gcc agc agt tct ttt gga tac aag aag cca agt ggt gta	1102
Arg Ser Thr Ala Ser Ser Phe Gly Tyr Lys Pro Ser Gly Val	
ggg gct tcc act atg att acc agc agc ggt gcc acc atc aca agc ggt	1150
Gly Ala Ser Thr Met Ile Thr Ser Ser Gly Ala Thr Ile Thr Ser Gly	
tca gct aca ctg ggg aaa atc ccc aaa tcc gct gcc att ggt ggg aag	1198
Ser Ala Thr Leu Gly Lys Ile Pro Lys Ser Ala Ala Ile Gly Gly Lys	
tcc aat gca gga agg aaa acc agc ctg gac ggg tcc cag aat caa gat	1246
Ser Asn Ala Gly Arg Lys Thr Ser Leu Asp Gly Ser Gln Asn Gln Asp	
gat gtt gtc ctg cac gtg agc tcg aag acc acc ctc cag tac cgt agt	1294
Asp Val Val Leu His Val Ser Ser Lys Thr Thr Leu Gln Tyr Arg Ser	
ttg ccc cgc cct tct aag tcc agc acc agc gga atc cct ggg aga ggt	1342
Leu Pro Arg Pro Ser Lys Ser Ser Thr Ser Gly Ile Pro Gly Arg Gly	

ggc cac agg tcg agc acc agc agc att gat tcc aat gtc agc agc aag Gly His Arg Ser Ser Thr Ser Ser Ile Asp Ser Asn Val Ser Ser Lys	1390
tca gct ggg gcc acc acc tcc aaa ctg aga gaa ccg act aag atc ggc Ser Ala Gly Ala Thr Thr Ser Lys Leu Arg Glu Pro Thr Lys Ile Gly	1438
tca ggg cgc tcg agt cca gtc act gtc aac caa aca gac aaa gag aag Ser Gly Arg Ser Ser Pro Val Thr Val Asn Gln Thr Asp Lys Glu Lys	1486
gag aaa gta gca gtg tca gat tca gag agc gtt tcc ttg tca ggt tcc Glu Lys Val Ala Val Ser Asp Ser Glu Ser Val Ser Leu Ser Gly Ser	1534
ccc aaa tcc agc ccc acc tct gcc agt gcc tgt ggg act caa ggg ctc Pro Lys Ser Ser Pro Thr Ser Ala Ser Ala Cys Gly Thr Gln Gly Leu	1582
aga cag cca ggg tcc aaa tat cca gat att gcc tcg ccc aca ttt cga Arg Gln Pro Gly Ser Lys Tyr Pro Asp Ile Ala Ser Pro Thr Phe Arg	1630
agg ttg ttc ggt gcc aag gca ggc ggc aaa tct gcc tcc gca cct aat Arg Leu Phe Gly Ala Lys Ala Gly Lys Ser Ala Ser Ala Pro Asn	1678
act gag ggg gcg aag tcc tcc tca gta gtg ctc agc cct agt acc tct Thr Glu Gly Ala Lys Ser Ser Val Val Leu Ser Pro Ser Thr Ser	1726
tta gcc cga caa ggc agt ctg gag tca ccg tcg tcc ggt acg gga agc Leu Ala Arg Gln Gly Ser Leu Glu Ser Pro Ser Gly Thr Gly Ser	1774
atg ggc agt gct ggt ggg ctg agt ggc agc agc agc cct ctc ttc aat Met Gly Ser Ala Gly Gly Leu Ser Gly Ser Ser Pro Leu Phe Asn	1822
aaa ccc tca gac cta act aca gat gtt ata agc tta agt cac tcc ttg Lys Pro Ser Asp Leu Thr Asp Val Ile Ser Leu Ser His Ser Leu	1870
gct tcc agc cca gcg tcg gtt cac tct ttc aca tcc ggt ggg ctt gtg Ala Ser Ser Pro Ala Ser Val His Ser Phe Thr Ser Gly Gly Leu Val	1918
tgg gct gcc aat ctg agc agt tcc tct gcc ggc agc aag gac act cca Trp Ala Ala Asn Leu Ser Ser Ala Gly Ser Lys Asp Thr Pro	1966
agt tac cag tcc atg act agt ctc cat acg agc tct gag tcc att gac Ser Tyr Gln Ser Met Thr Ser Leu His Thr Ser Ser Glu Ser Ile Asp	2014

ctg ccc ctc agc cat cat ggc tcc ctg tct gga ctg acc aca ggc act	2062
Leu Pro Leu Ser His His Gly Ser Leu Ser Gly Leu Thr Thr Gly Thr	
cac gag gtg cag agc ctc atg aga acg ggt agt gtg aga tct act	2110
His Glu Val Gln Ser Leu Leu Met Arg Thr Gly Ser Val Arg Ser Thr	
ctc tca gaa aga tac acc cca tca tct cgg cag gcc aac caa gaa gaa	2158
Leu Ser Glu Arg Tyr Thr Pro Ser Ser Arg Gln Ala Asn Gln Glu Glu	
ggc aaa gag tgg ctg cga tcg cat tcc act ggc ggg ctg cag gat act	2206
Gly Lys Glu Trp Leu Arg Ser His Ser Thr Gly Gly Leu Gln Asp Thr	
ggc aac cag tct ccc ttg gtc tcc cct tct gcc atg tca tcg tca gcc	2254
Gly Asn Gln Ser Pro Leu Val Ser Pro Ser Ala Met Ser Ser Ala	
acc gga aaa tat cac ttt tcc aac ttg gtg agt ccc acc aac ctc tcc	2302
Thr Gly Lys Tyr His Phe Ser Asn Leu Val Ser Pro Thr Asn Leu Ser	
cag ttt aac ctg cct gca ccc agt atg atg cgc tcc agc agt atc ccc	2350
Gln Phe Asn Leu Pro Ala Pro Ser Met Met Arg Ser Ser Ser Ile Pro	
gcc cag gac tcc tcc ttc gac ctc tat gat gat gcc cag ctt tgc ggt	2398
Ala Gln Asp Ser Ser Phe Asp Leu Tyr Asp Ala Gln Leu Cys Gly	
agt gca act tcc ctg gag gaa agg cca cgg gcc gtt agc cac tcc ggc	2446
Ser Ala Thr Ser Leu Glu Arg Pro Arg Ala Val Ser His Ser Gly	
tca ttc aga gac agc atg gag gaa gtt cat ggc tct tca ctg tca ttg	2494
Ser Phe Arg Asp Ser Met Glu Glu Val His Gly Ser Ser Leu Ser Leu	
gtc tcc agc aca tca tcc ctt tac tct acg gct gaa gag aag gct cat	2542
Val Ser Ser Thr Ser Leu Tyr Ser Thr Ala Glu Glu Lys Ala His	
tca gag caa atc cat aag cta cgg aga gaa ctg gtt gcc tcc cag gag	2590
Ser Glu Gln Ile Lys Leu Arg Arg Glu Leu Val Ala Ser Gln Glu	
aaa gtc gct acc ctc acg tct cag ctg tca gca aat gct cac ctt gta	2638
Lys Val Ala Thr Leu Thr Ser Gln Leu Ser Ala Asn Ala His Leu Val	
gca gct ttt gaa aag agt tta ggg aat atg act ggc cgt ttg caa agt	2686
Ala Ala Phe Glu Lys Ser Leu Gly Asn Met Thr Gly Arg Leu Gln Ser	
cta acc atg aca gcg gaa caa aag gaa tct gag ctt atc gaa ctg cgg	2734

Leu Thr Met Thr Ala Glu Gln Lys Glu Ser Glu Leu Ile Glu Leu Arg  
 gaa acc att gaa atg ttg aag gcc cag aac tct gct gcc caa gca gcc 2782  
 Glu Thr Ile Glu Met Leu Lys Ala Gln Asn Ser Ala Ala Gln Ala Ala  
  
 att cag gga gca ctg aat ggc cca gac cac cct ccc aaa gat ctc cgc 2830  
 Ile Gln Gly Ala Leu Asn Gly Pro Asp His Pro Pro Lys Asp Leu Arg  
  
 atc aga aga cag cac tcc tct gaa agt gtt tct agt atc aac agc gca 2878  
 Ile Arg Arg Gln His Ser Ser Glu Ser Val Ser Ser Ile Asn Ser Ala  
  
 acg agc cat tcc agc att ggc agt ggt aat gat gct gac tcc aag aaa 2926  
 Thr Ser His Ser Ser Ile Gly Ser Gly Asn Asp Ala Asp Ser Lys Lys  
  
 <210> 20  
 <211> 975  
 <212> PRT  
 <213> Mouse  
  
 <400> 20

Ser	His	Ser	Thr	Leu	Glu	Thr	Thr	Phe	Asp	Thr	Thr	Val	Thr	Thr	Glu
1				5					10				15		

Val	Asn	Gly	Arg	Ala	Ile	Pro	Asn	Leu	Thr	Ser	Arg	Pro	Ser	Pro	Met
				20				25				30			

Thr	Trp	Arg	Leu	Gly	Gln	Ala	Cys	Pro	Arg	Leu	Gln	Ala	Gly	Asp	Ala
					35		40				45				

Pro	Ser	Met	Gly	Ala	Gly	Tyr	Ser	Arg	Ser	Gly	Thr	Ser	Arg	Phe	Ile
						50	55				60				

His	Thr	Asp	Pro	Ser	Arg	Phe	Met	Tyr	Thr	Thr	Pro	Leu	Arg	Arg	Ala
							65	70			75		80		

Ala	Val	Ser	Arg	Leu	Gly	Asn	Met	Ser	Gln	Ile	Asp	Met	Ser	Glu	Lys
						85			90				95		

Ala	Ser	Ser	Asp	Leu	Asp	Val	Ser	Ser	Glu	Val	Asp	Val	Gly	Gly	Tyr
						100		105				110			

Met	Ser	Asp	Gly	Asp	Ile	Leu	Gly	Lys	Ser	Leu	Arg	Ala	Asp	Asp	Ile
								115	120			125			

Asn	Ser	Gly	Tyr	Met	Thr	Asp	Gly	Gly	Leu	Asn	Leu	Tyr	Thr	Arg	Ser
								130	135			140			

Leu	Asn	Arg	Val	Pro	Asp	Thr	Ala	Thr	Ser	Arg	Asp	Val	Ile	Gln	Arg
								145	150			155		160	

Gly Val His Asp Val Thr Val Asp Ala Asp Ser Trp Asp Asp Ser Ser  
165 170 175

Ser Val Ser Ser Gly Leu Ser Asp Thr Leu Asp Asn Ile Ser Thr Asp  
180 185 190

Asp Leu Asn Thr Thr Ser Ser Ile Ser Ser Tyr Ser Asn Ile Thr Val  
195 200 205

Pro Ser Arg Lys Asn Thr Gln Leu Lys Thr Asp Ala Glu Lys Arg Ser  
210 215 220

Thr Thr Asp Glu Thr Trp Asp Ser Pro Glu Glu Leu Lys Lys Ala Glu  
225 230 235 240

Gly Asp Cys Asp Ser His Gly Asp Gly Ala Ala Lys Trp Lys Gly Ala  
245 250 255

Thr Ser Gly Leu Ala Glu Asp Ser Glu Lys Thr Gly Gln Lys Ala Ser  
260 265 270

Leu Ser Val Ser Gln Thr Gly Ser Trp Arg Arg Gly Met Ser Ala Gln  
275 280 285

Gly Gly Thr Pro Ala Thr Ala Arg Gln Lys Thr Ser Thr Ser Ala Leu  
290 295 300

Lys Thr Pro Gly Lys Thr Asp Asp Ala Lys Ala Ser Glu Lys Gly Lys  
305 310 315 320

Thr Pro Leu Lys Gly Ser Ser Leu Gln Arg Ser Pro Ser Asp Ala Gly  
325 330 335

Lys Ser Ser Gly Asp Glu Gly Lys Lys Pro Pro Ser Gly Ile Gly Arg  
340 345 350

Ser Thr Ala Ser Ser Ser Phe Gly Tyr Lys Lys Pro Ser Gly Val Gly  
355 360 365

Ala Ser Thr Met Ile Thr Ser Ser Gly Ala Thr Ile Thr Ser Gly Ser  
370 375 380

Ala Thr Leu Gly Lys Ile Pro Lys Ser Ala Ala Ile Gly Gly Lys Ser  
385 390 395 400

Asn Ala Gly Arg Lys Thr Ser Leu Asp Gly Ser Gln Asn Gln Asp Asp  
405 410 415

Val Val Leu His Val Ser Ser Lys Thr Thr Leu Gln Tyr Arg Ser Leu  
420 425 430

Pro Arg Pro Ser Lys Ser Ser Thr Ser Gly Ile Pro Gly Arg Gly Gly  
435 440 445

His Arg Ser Ser Thr Ser Ser Ile Asp Ser Asn Val Ser Ser Lys Ser

450	455	460
Ala Gly Ala Thr Thr Ser Lys Leu Arg Glu Pro Thr Lys Ile Gly Ser		
465	470	475
Gly Arg Ser Ser Pro Val Thr Val Asn Gln Thr Asp Lys Glu Lys Glu		
485	490	495
Lys Val Ala Val Ser Asp Ser Glu Ser Val Ser Leu Ser Gly Ser Pro		
500	505	510
Lys Ser Ser Pro Thr Ser Ala Ser Ala Cys Gly Thr Gln Gly Leu Arg		
515	520	525
Gln Pro Gly Ser Lys Tyr Pro Asp Ile Ala Ser Pro Thr Phe Arg Arg		
530	535	540
Leu Phe Gly Ala Lys Ala Gly Gly Lys Ser Ala Ser Ala Pro Asn Thr		
545	550	555
Glu Gly Ala Lys Ser Ser Ser Val Val Leu Ser Pro Ser Thr Ser Leu		
565	570	575
Ala Arg Gln Gly Ser Leu Glu Ser Pro Ser Ser Gly Thr Gly Ser Met		
580	585	590
Gly Ser Ala Gly Gly Leu Ser Gly Ser Ser Ser Pro Leu Phe Asn Lys		
595	600	605
Pro Ser Asp Leu Thr Thr Asp Val Ile Ser Leu Ser His Ser Leu Ala		
610	615	620
Ser Ser Pro Ala Ser Val His Ser Phe Thr Ser Gly Gly Leu Val Trp		
625	630	640
Ala Ala Asn Leu Ser Ser Ser Ala Gly Ser Lys Asp Thr Pro Ser		
645	650	655
Tyr Gln Ser Met Thr Ser Leu His Thr Ser Ser Glu Ser Ile Asp Leu		
660	665	670
Pro Leu Ser His His Gly Ser Leu Ser Gly Leu Thr Thr Gly Thr His		
675	680	685
Glu Val Gln Ser Leu Leu Met Arg Thr Gly Ser Val Arg Ser Thr Leu		
690	695	700
Ser Glu Arg Tyr Thr Pro Ser Ser Arg Gln Ala Asn Gln Glu Glu Gly		
705	710	715
Lys Glu Trp Leu Arg Ser His Ser Thr Gly Gly Leu Gln Asp Thr Gly		
725	730	735
Asn Gln Ser Pro Leu Val Ser Pro Ser Ala Met Ser Ser Ala Thr		
740	745	750

Gly	Lys	Tyr	His	Phe	Ser	Asn	Leu	Val	Ser	Pro	Thr	Asn	Leu	Ser	Gln
755															765
Phe	Asn	Leu	Pro	Ala	Pro	Ser	Met	Met	Arg	Ser	Ser	Ser	Ile	Pro	Ala
770															780
Gln	Asp	Ser	Ser	Phe	Asp	Leu	Tyr	Asp	Asp	Ala	Gln	Leu	Cys	Gly	Ser
785															800
Ala	Thr	Ser	Leu	Glu	Glu	Arg	Pro	Arg	Ala	Val	Ser	His	Ser	Gly	Ser
															815
805															
Phe	Arg	Asp	Ser	Met	Glu	Glu	Val	His	Gly	Ser	Ser	Leu	Ser	Leu	Val
															830
Ser	Ser	Thr	Ser	Ser	Leu	Tyr	Ser	Thr	Ala	Glu	Glu	Lys	Ala	His	Ser
															845
Glu	Gln	Ile	His	Lys	Leu	Arg	Arg	Glu	Leu	Val	Ala	Ser	Gln	Glu	Lys
															860
850															
Val	Ala	Thr	Leu	Thr	Ser	Gln	Leu	Ser	Ala	Asn	Ala	His	Leu	Val	Ala
															880
865															
Ala	Phe	Glu	Lys	Ser	Leu	Gly	Asn	Met	Thr	Gly	Arg	Leu	Gln	Ser	Leu
															895
885															
Thr	Met	Thr	Ala	Glu	Gln	Lys	Glu	Ser	Glu	Leu	Ile	Glu	Leu	Arg	Glu
															910
900															
Thr	Ile	Glu	Met	Leu	Lys	Ala	Gln	Asn	Ser	Ala	Ala	Gln	Ala	Ala	Ile
															925
915															
Gln	Gly	Ala	Leu	Asn	Gly	Pro	Asp	His	Pro	Pro	Lys	Asp	Leu	Arg	Ile
															940
930															
Arg	Arg	Gln	His	Ser	Ser	Glu	Ser	Val	Ser	Ser	Ile	Asn	Ser	Ala	Thr
															960
945															
Ser	His	Ser	Ser	Ile	Gly	Ser	Gly	Asn	Asp	Ala	Asp	Ser	Lys	Lys	
															975
965															

<210> 21  
<211> 2583  
<212> DNA  
<213> mouse

<400> 21

gggatgaagg	aaaaaagcca	ccgtcaggca	ttggaagatc	gacagccagc	agttcttttg	60
gataacaagaa	gccaaagtgg	gtaggggctt	ccactatgtat	taccaggcgc	ggtgccacca	120
tcacaagcgg	ttcagctaca	ctggggaaaa	tcccccaaatc	cgctgccatt	ggtgggaagt	180

ccaatgcagg aaggaaaacc agcctggacg ggtcccagaa tcaagatgat gttgtcctgc	240
acgtgagctc gaagaccacc ctccagtacc gtagttgcc cccgccttct aagtccagca	300
ccagcgaaat ccctgggaga ggtggccaca ggtcgagcac cagcagcatt gattccaatg	360
tcagcagcaa gtcagctggg gccaccaccc ccaaactgag agaaccgact aagatcggt	420
cagggcgctc gagtccagtc actgtcaacc aaacagacaa agagaaggag aaagtagcag	480
tgtcagattc agagagcggtt tccttgcag gttccccaa atccagcccc acctctgcca	540
gtgcctgtgg gactcaaggg ctcagacagc cagggccaa atatccagat attgcctcgc	600
ccacatttcg aaggtaaggg tatgtaaaga gatgttgaa aacataaaa ggtagtatat	660
agcatgtatt tattctgtac gaaactattt tcatgtattc taaatattct aagattctgt	720
atcttatact tgtctaaaat atagtgattt tattttgctg attgcacctg ttgctagtgt	780
aaaagcattt ctcatttaga gagtggttag ccttcagct atacagccag tgtgacacta	840
aaatacagat accacttgcataa gcgggcataa aaccacatga ctgactatcc atagaaataa	900
agtatgtact tgtaaagata ttttagtgatt tccacccctc ctttccagaa ttaaaaaaaaag	960
caaattgcat agatctttat aaacacattt acttctagtg tatgttatct tggtactct	1020
taatgaaatg gcagttatga atatagatga tatattctt ctaacagttt ataagagacc	1080
aatttataca gtaccagatc ttaacatagt aacaataaca gcaacaaaaa caacccaaaa	1140
agctatcaa gtatggctg attgcagaat ttgaaaacat ttacatgtt gacataggac	1200
aagaactcag gagtgaggtg acttttata agtcttcatc aatgtccctt tacaggaacc	1260
aggaagcata tctgatatat gtgtcaggat tatcacttta ttaattatgt gaaattctgt	1320
ttagaaatct acctgatttt aaatacttta atatagtagg ggtcaaaatt agttaatgag	1380
ttaagacaag ttgttaaata atcctggctc tgccccctca tcttcaaaaat gatagagtt	1440
aatttatac cttttgttaa atattcagg tttgtgtttt ttctcttgcatt aactttgatc	1500
tcttagaaga gtcttgaaga atttacatta agtaatcttta gaaacataac tatttgagaa	1560
acagtagtca aattttgtca tttagaagtat taactctgaa gaatgattt aagtgacagt	1620
tcttagaaag aataaaattat agctttagc aagagtaaat attttcactg cttgtgtgag	1680
agccaaagagc gccctcttgt ggcccattac ctatgaaaca atttctcata ttcccccctag	1740
aaatcttcca ctgcaggaaa taatggattt cattgcctct gaatttagtaa ccatctgcc	1800
atttcttcat accattttat ttccataactt gcataaattt gattatgtca tctgtttcat	1860
ttacaaaact aaaatgtttt ctgagctaaa ctccagtagc taacttagta caaatggat	1920

ttttaaatca ctgctataag tatatatatt tgaatagctc tggcaacgga cgaaaaatccc 1980  
tatggtcttt ccatggaaag atacaaaccca atccataagt tgtccagcaa tatccaatat 2040  
ttccagccca gccagtcagg cctcttaaac attacctac atatgtAAC ctttccttaa 2100  
atgtccccct tagacaatct atttttaaa aagatgaaaa tccatTTAA catcatatat 2160  
cgaatgcgta gaagttgttt cattataatg gttctgcaga tagtaatgc caaaacggcc 2220  
aaaatatttg atcaactagaa gcttaaaagt caagtacaat catgttact ttTTTCCAA 2280  
ggtgggttca ctgctgccc ccttgggttcc aggccagtgc ttacttaaga tatcgtaagt 2340  
gattttttt taattttaa ttttttagta gttggtaat caaaagccag tcatgtcacc 2400  
ttcaggaaca tagaggctgg acgtgcttgg cagctcacga ctccaaagca cactggctc 2460  
tgtggactga aacccttagga aacgtggatg tgagtctttt ggaacaactc aagttgttat 2520  
ttgttttct ttttaggtgt tcggtgccaa ggcaggcggc aaatctgcct ccgcacctaa 2580  
tac 2583

<210> 22  
<211> 37  
<212> DNA  
<213> Homo sapiens  
  
<220>  
<221> CDS  
<222> (2)..(34)  
  
<400> 22

c cgc ggg gct tcc atc ctt cct ttg act gat ttt taa 37  
Arg Gly Ala Ser Ile Leu Pro Leu Thr Asp Phe \*

<210> 23  
<211> 6768  
<212> DNA  
<213> Homo sapiens  
  
<220>  
<221> CDS  
<222> (1)..(5928)  
  
<400> 23

att tta att tgt att ttc ccc gcc gcc ccg ccc ctt ttc ctc cga ccc 48

Ile Leu Ile Cys Ile Phe Pro Ala Ala Pro Pro Leu Phe Leu Arg Pro	
 cgc cct atc gct ccc cggttccc tgc tct ttc ctt ttt ccc ggc ttc Arg Pro Ile Ala Pro Arg Leu Pro Cys Ser Phe Leu Phe Pro Gly Phe	96
 ctt cct cgc gtt tct ttc ccc tgc gcc ctc ggc ttg cct ctc tcc ctc Leu Pro Arg Val Ser Phe Pro Cys Ala Leu Gly Leu Pro Leu Ser Leu	144
 ctc cct cgc tct ctc ccc ctt ctc tcc cct tct tcc tcg gtt tct tcc Leu Pro Arg Ser Leu Pro Leu Leu Ser Pro Ser Ser Val Ser Ser	192
 gtc ctc tct ctc ccc ctc ctc ccc cgc ctc ctc ctg cgc tcc Val Leu Ser Leu Pro Leu Leu Pro Arg Leu Leu Leu Arg Ser	240
 cgc ccc ctg ccc cct ccc ccc gtg cct gca gac gcg cgg atc gtc cat Arg Pro Leu Pro Pro Val Pro Ala Asp Ala Arg Ile Val His	288
 gcg ctc ctc gcg ggc aga atg ctg ggc agc agc gtc aag agc gtg cag Ala Leu Leu Ala Gly Arg Met Leu Gly Ser Ser Val Lys Ser Val Gln	336
 ccc gag gtg gag ctg agc agc ggc ggc gac gag ggc gcg gac gaa Pro Glu Val Glu Leu Ser Ser Gly Gly Asp Glu Gly Ala Asp Glu	384
 ccg cgg ggc gcc ggc agg aag gcg gca gcg gcg gac ggc aga ggc atg Pro Arg Gly Ala Gly Arg Lys Ala Ala Ala Asp Gly Arg Gly Met	432
 ctg ccc aag cgc gcc aag gcg ccc ggc ggc ggc ggc atg gcc aag Leu Pro Lys Arg Ala Lys Ala Pro Gly Gly Gly Met Ala Lys	480
 gcc agc gcg gct gag ctg aag gtc ttc aag tcc ggc agc gtg gac agc Ala Ser Ala Ala Glu Leu Lys Val Phe Lys Ser Gly Ser Val Asp Ser	528
 cgt gtc ccc ggc ggg ccg ccc gcc tcc aac ctg cgc aag cag aag tca Arg Val Pro Gly Pro Pro Ala Ser Asn Leu Arg Lys Gln Lys Ser	576
 ctc acc aac ctc tct ttt ctc acg gac tcc gag aaa aag ctg cag ctt Leu Thr Asn Leu Ser Phe Leu Thr Asp Ser Glu Lys Lys Leu Gln Leu	624
 tat gag ccc gaa tgg agc gac gat atg gcc aag gcg ccc aaa ggc tta Tyr Glu Pro Glu Trp Ser Asp Asp Met Ala Lys Ala Pro Lys Gly Leu	672
 ggc aag gtg ggg tcc aag ggc cgt gaa gct ccg ctg atg tcc aag acg Gly Lys Val Gly Ser Lys Gly Arg Glu Ala Pro Leu Met Ser Lys Thr	720

ctg tcc aag tcg gag cac tcg ctc ttc cag gcc aag ggc agc ccg gcg	768
Leu Ser Lys Ser Glu His Ser Leu Phe Gln Ala Lys Gly Ser Pro Ala	
ggc ggc gcc aag acc ccc ctg gct ccg ctc gcg ccc aac ctg gga aag	816
Gly Gly Ala Lys Thr Pro Leu Ala Pro Leu Ala Pro Asn Leu Gly Lys	
ccg agc cgg atc cct cga gga ccc tat gcg gag gtc aag ccg ctc agc	864
Pro Ser Arg Ile Pro Arg Gly Pro Tyr Ala Glu Val Lys Pro Leu Ser	
aag gcg cct gaa gcg gcc gtg agc gaa gat ggc aaa tcg gac gac gag	912
Lys Ala Pro Glu Ala Ala Val Ser Glu Asp Gly Lys Ser Asp Asp Glu	
ctg ctc tcc agc aag gcc aag gcg caa aag agc tct ggg cct gtc ccc	960
Leu Leu Ser Ser Lys Ala Lys Ala Gln Lys Ser Ser Gly Pro Val Pro	
tct gcc aag ggc cag gag gag cgc gcc ttc ctc aag gtg gac ccc gag	1008
Ser Ala Lys Gly Gln Glu Arg Ala Phe Leu Lys Val Asp Pro Glu	
ctg gtg gtg acc gtg ctg gga gac ctg gag cag ctg ctc ttc agc cag	1056
Leu Val Val Thr Val Leu Gly Asp Leu Glu Gln Leu Leu Phe Ser Gln	
atg ctg gac cca gag tcc cag aga aag agg aca gtg cag aat gtc ctg	1104
Met Leu Asp Pro Glu Ser Gln Arg Lys Arg Thr Val Gln Asn Val Leu	
gat ctc cgg cag aac ctg gaa gag acc atg tcc agc ctg cga ggg tcc	1152
Asp Leu Arg Gln Asn Leu Glu Glu Thr Met Ser Ser Leu Arg Gly Ser	
cag gtg act cac agc tcc ctg gag atg acc tgc tac gac agc gat gat	1200
Gln Val Thr His Ser Ser Leu Glu Met Thr Cys Tyr Asp Ser Asp Asp	
gcc aac cca cgc agc gtg tcc agc ctc tcc aac cgc tcg tac cct ctg	1248
Ala Asn Pro Arg Ser Val Ser Ser Leu Ser Asn Arg Ser Tyr Pro Leu	
tca tgg cgc tat ggc cag tcc agt ccg cgg ctg cag gct ggt gac gcg	1296
Ser Trp Arg Tyr Gln Ser Ser Pro Arg Leu Gln Ala Gly Asp Ala	
ccc tct gtg ggt ggg agc tgc cgc tcg gag ggg acg ccc gcc tgg tac	1344
Pro Ser Val Gly Gly Ser Cys Arg Ser Glu Gly Thr Pro Ala Trp Tyr	
atg cac ggc gaa cgg gcc cac tac tcc cac acc atg ccc atg cgc agc	1392
Met His Gly Glu Arg Ala His Tyr Ser His Thr Met Pro Met Arg Ser	

ccc agc aag ctc agc cat atc tcc cgc ctg gag ctg gtc gaa tcc ctg Pro Ser Lys Leu Ser His Ile Ser Arg Leu Glu Leu Val Glu Ser Leu	1440
gac tcg gat gag gtg gac ctc aag tcc ggc tac atg agc gac agt gac Asp Ser Asp Glu Val Asp Leu Lys Ser Gly Tyr Met Ser Asp Ser Asp	1488
ctc atg ggc aag acc atg acg gag gat gat gac atc act acc ggc tgg Leu Met Gly Lys Thr Met Thr Glu Asp Asp Asp Ile Thr Thr Gly Trp	1536
gat gaa agc agc tcc atc agt agt gga ctc agc gat gcc tca gac aat Asp Glu Ser Ser Ser Ile Ser Ser Gly Leu Ser Asp Ala Ser Asp Asn	1584
ctc agt tca gaa gaa ttc aat gcc agc tcc tca aac tcc ctc cca Leu Ser Ser Glu Glu Phe Asn Ala Ser Ser Leu Asn Ser Leu Pro	1632
agt act ccc act gct tct cgc agg aac tca aca ata gtg cta cgc aca Ser Thr Pro Thr Ala Ser Arg Arg Asn Ser Thr Ile Val Leu Arg Thr	1680
gac tca gag aag cgc tca ctg gca gaa agt ggg ctg agc tgg ttt agt Asp Ser Glu Lys Arg Ser Leu Ala Glu Ser Gly Leu Ser Trp Phe Ser	1728
gaa tca gag gag aaa gcc cct aaa aaa ctg gag tac gac agt ggt agc Glu Ser Glu Glu Lys Ala Pro Lys Lys Leu Glu Tyr Asp Ser Gly Ser	1776
ctg aag atg gaa cct ggg act tct aag tgg cgg agg gag cgg cct gag Leu Lys Met Glu Pro Gly Thr Ser Lys Trp Arg Arg Glu Arg Pro Glu	1824
agc tgt gat gat tca tcc aag ggt gga gaa ctg aaa aag ccc atc agc Ser Cys Asp Asp Ser Ser Lys Gly Glu Leu Lys Lys Pro Ile Ser	1872
ctg ggc cac cct ggt tcc ctg aag aag ggc aag acc cca cct gtg gct Leu Gly His Pro Gly Ser Leu Lys Lys Gly Lys Thr Pro Pro Val Ala	1920
gta act tcc ccc atc act cac aca gcc cag agt gcc ctc aaa gtc gca Val Thr Ser Pro Ile Thr His Thr Ala Gln Ser Ala Leu Lys Val Ala	1968
ggc aaa cct gag ggc aaa gct aca gac aag ggt aag ctt gca gtg aag Gly Lys Pro Glu Gly Lys Ala Thr Asp Lys Gly Lys Leu Ala Val Lys	2016
aat act ggg ctc caa cgc tcc tcc tct gat gct ggt cgg gac cgc ctg Asn Thr Gly Leu Gln Arg Ser Ser Asp Ala Gly Arg Asp Arg Leu	2064

agt gat gct aag aag ccc ccc tcg ggc att gct cgc ccc tcc act tcg	2112
Ser Asp Ala Lys Lys Pro Pro Ser Gly Ile Ala Arg Pro Ser Thr Ser	
gga tcc ttt ggc tac aag aag cct cct gcc aca ggc aca gcc act	2160
Gly Ser Phe Gly Tyr Lys Pro Pro Ala Thr Gly Thr Ala Thr	
gtc atg caa act ggt ggt tca gcc act ctc agc aag atc cag aag tcc	2208
Val Met Gln Thr Gly Ser Ala Thr Leu Ser Lys Ile Gln Lys Ser	
tca ggc atc cct gtc aag cca gta aat ggg cgc aag act agc tta gat	2256
Ser Gly Ile Pro Val Lys Pro Val Asn Gly Arg Lys Thr Ser Leu Asp	
gtt tcc aac agt gca gag cca gga ttc ctg gct cct gga gcc cgt tct	2304
Val Ser Asn Ser Ala Glu Pro Gly Phe Leu Ala Pro Gly Ala Arg Ser	
aac atc cag tac cgc agc ctg ccc cg <sup>g</sup> cca gcc aag tca agt tct atg	2352
Asn Ile Gln Tyr Arg Ser Leu Pro Arg Pro Ala Lys Ser Ser Ser Met	
agc gtg acc ggc ggg cg <sup>g</sup> ggt gga cct cgc cct gtg agc agc agc att	2400
Ser Val Thr Gly Arg Gly Pro Arg Pro Val Ser Ser Ser Ile	
gac ccc agt ctc ctc agc acc aag cag gga ggc ctt acg cct tcc aga	2448
Asp Pro Ser Leu Leu Ser Thr Lys Gln Gly Leu Thr Pro Ser Arg	
ctg aag gag cct acc aag gta gcc agt ggg cg <sup>g</sup> acc act cca gcc cct	2496
Leu Lys Glu Pro Thr Lys Val Ala Ser Gly Arg Thr Thr Pro Ala Pro	
gtc aat cag aca gat cg <sup>g</sup> gaa aag gag aag gcc aaa gcc aag gca gtg	2544
Val Asn Gln Thr Asp Arg Glu Lys Ala Lys Ala Lys Ala Val	
gcc ttg gac tca gac aac atc tcc ttg aag agt att ggc tcc cca gaa	2592
Ala Leu Asp Ser Asp Asn Ile Ser Leu Lys Ser Ile Gly Ser Pro Glu	
agt act ccc aag aac caa gca agc cac ccc aca gcc acc aag ctg gca	2640
Ser Thr Pro Lys Asn Gln Ala Ser His Pro Thr Ala Thr Lys Leu Ala	
gag ctg cca cca acc cct ctc agg gcc aca gc <sup>g</sup> aag agc ttt gtc aaa	2688
Glu Leu Pro Pro Thr Pro Leu Arg Ala Thr Ala Lys Ser Phe Val Lys	
cca ccc tca cta gcc aat ctt gac aag gtc aac tcc aac agt ctg gat	2736
Pro Pro Ser Leu Ala Asn Leu Asp Val Asn Ser Asn Ser Leu Asp	
cta cca tca tcc agt gat acc acc cat gct tca aag gtc cca gat ctg	2784
Leu Pro Ser Ser Ser Asp Thr Thr His Ala Ser Lys Val Pro Asp Leu	

cat gct aca agc tca gca tct ggg ggc cct ctc cct tcc tgc ttc acc	2832
His Ala Thr Ser Ser Ala Ser Gly Gly Pro Leu Pro Ser Cys Phe Thr	
ccc agt ccg gca ccc atc ctc aat att aac tca gcc agc ttc tcc cag	2880
Pro Ser Pro Ala Pro Ile Leu Asn Ile Asn Ser Ala Ser Phe Ser Gln	
ggc ctg gag cta atg agt ggt ttc agt gtg cca aaa gag acc cgc atg	2928
Gly Leu Glu Leu Met Ser Gly Phe Ser Val Pro Lys Glu Thr Arg Met	
tac ccc aaa ctc tca ggc ctg cac agg agc atg gag tcc ctc cag atg	2976
Tyr Pro Lys Leu Ser Gly Leu His Arg Ser Met Glu Ser Leu Gln Met	
cca atg agc ctc ccc agt gcc ttc ccc agc agt act ccc gtc ccc acc	3024
Pro Met Ser Leu Pro Ser Ala Phe Pro Ser Ser Thr Pro Val Pro Thr	
cca cct gct ccc cct gct ccc aca gaa gaa gag acg gaa gag ctg	3072
Pro Pro Ala Pro Pro Ala Ala Pro Thr Glu Glu Thr Glu Glu Leu	
act tgg agt gga agc ccc aga gct ggg caa ctg gac agt aat cag cgg	3120
Thr Trp Ser Gly Ser Pro Arg Ala Gly Gln Leu Asp Ser Asn Gln Arg	
gat cgg aac act ctt ccc aag aaa ggg ctc agg tac cag ctt cag tcc	3168
Asp Arg Asn Thr Leu Pro Lys Lys Gly Leu Arg Tyr Gln Leu Gln Ser	
cag gag gag acc aag gag agg cga cat tcc cat acc att ggt ggg ctg	3216
Gln Glu Glu Thr Lys Glu Arg Arg His Ser His Thr Ile Gly Gly Leu	
cct gaa tcc gat gac cag tca gag ctg cct tct ccc cct gca ctt ccc	3264
Pro Glu Ser Asp Asp Gln Ser Glu Leu Pro Ser Pro Pro Ala Leu Pro	
atg tct ctg agt gca aag ggc caa ctt acc aac ata gtg agt ccc act	3312
Met Ser Leu Ser Ala Lys Gly Gln Leu Thr Asn Ile Val Ser Pro Thr	
gcg gcc acc acg cca aga atc acc cgc tcc aac agc atc ccc acc cac	3360
Ala Ala Thr Thr Pro Arg Ile Thr Arg Ser Asn Ser Ile Pro Thr His	
gag gcg gcc ttc gag ctg tac agc ggc tcc caa atg ggg agc acc ctg	3408
Glu Ala Ala Phe Glu Leu Tyr Ser Gly Ser Gln Met Gly Ser Thr Leu	
tcc ctg gcc gag aga ccc aag gga atg att cgg tca gga tcc ttc cga	3456
Ser Leu Ala Glu Arg Pro Lys Gly Met Ile Arg Ser Gly Ser Phe Arg	
gac ccc acg gac gat gtt cac ggc tca gtg ctg tcc ctg gcc tcc agt	3504
Asp Pro Thr Asp Asp Val His Gly Ser Val Leu Ser Leu Ala Ser Ser	
gcc tcc tcc acc tac tcc tca gct gag gag agg atg caa tct gag caa	3552
Ala Ser Ser Thr Tyr Ser Ala Glu Glu Arg Met Gln Ser Glu Gln	
atc cgg aag ctt cgt agg gaa ctg gaa tca tcc cag gaa aaa gtg gcc	3600
Ile Arg Lys Leu Arg Arg Glu Leu Ser Ser Gln Glu Lys Val Ala	
acc ttg acg tct cag ctt tct gcc aat gct aat ctg gtg gct gct ttt	3648
Thr Leu Thr Ser Gln Leu Ser Ala Asn Ala Asn Leu Val Ala Ala Phe	
gag cag agc ctg gtg aat atg aca tcc cgc ctg cga cac ctg gca gag	3696
Glu Gln Ser Leu Val Asn Met Thr Ser Arg Leu Arg His Leu Ala Glu	

acg gcc gag gag aag gac act gag ctg ctg gat ttg cga gaa acc ata	3744
Thr Ala Glu Glu Lys Asp Thr Glu Leu Leu Asp Leu Arg Glu Thr Ile	
gac ttt ctg aag aaa aag aac tct gag gcc cag gca gtc att cag gga	3792
Asp Phe Leu Lys Lys Asn Ser Glu Ala Gln Ala Val Ile Gln Gly	
gcc ctt aat gcc tca gaa acc aca ccc aaa gaa ctt cg <sup>g</sup> atc aag aga	3840
Ala Leu Asn Ala Ser Glu Thr Thr Pro Lys Glu Leu Arg Ile Lys Arg	
caa aac tcc tca gat agc atc tca agc ctc aac agc atc act agc cat	3888
Gln Asn Ser Ser Asp Ser Ile Ser Ser Leu Asn Ser Ile Thr Ser His	
tcc agc atc ggc agc agc aag gat gct gat gcg aaa aag aag aaa aaa	3936
Ser Ser Ile Gly Ser Ser Lys Asp Ala Asp Ala Lys Lys Lys Lys	
aag agt tgg ctt cga agt tcc ttc aac aaa gcg ttc agt ata aaa aag	3984
Lys Ser Trp Leu Arg Ser Phe Asn Lys Ala Phe Ser Ile Lys Lys	
ggg ccc aag tca gct tcc tca tac tcg gat ata gag gag att gct aca	4032
Gly Pro Lys Ser Ala Ser Ser Tyr Ser Asp Ile Glu Glu Ile Ala Thr	
ccc gac tct tca gcc ccc tca tcc ccc aaa cta cag cat ggt tct aca	4080
Pro Asp Ser Ser Ala Pro Ser Ser Pro Lys Leu Gln His Gly Ser Thr	
gag act gct tca ccc tcc atc aag tcc tcc acc tcg tcc gtg ggc	4128
Glu Thr Ala Ser Pro Ser Ile Lys Ser Ser Thr Ser Ser Val Gly	
act gat gtc acc gag ggc cct gct cac cca gcc ccc cac act agg ctg	4176
Thr Asp Val Thr Glu Gly Pro Ala His Pro Ala Pro His Thr Arg Leu	
ttc cat gca aat gag gag gag cca gag aag gag gta tcg gag	4224
Phe His Ala Asn Glu Glu Pro Glu Lys Lys Glu Val Ser Glu	
ctg cgc tct gag cta tgg gag aag gaa atg aag ctt aca gac atc cgc	4272
Leu Arg Ser Glu Leu Trp Glu Lys Glu Met Lys Leu Thr Asp Ile Arg	
ttg gag gcc ctc aac tct gcc cac caa ctg gat cag ctt cg <sup>g</sup> gag acc	4320
Leu Glu Ala Leu Asn Ser Ala His Gln Leu Asp Gln Leu Arg Glu Thr	
atg cac aac atg cag ttg gag gtg gac ctg ctg gaa gca gag aat gac	4368
Met His Asn Met Gln Leu Glu Val Asp Leu Leu Glu Ala Glu Asn Asp	
cga ctg aag gta gcc cca ggc ccc tca tca ggc tcc act cca ggg cag	4416
Arg Leu Lys Val Ala Pro Gly Pro Ser Ser Gly Ser Thr Pro Gly Gln	
gtc cct gga tca tct gca tta tct tcc cca cgc cgc tcc cta ggc ctg	4464
Val Pro Gly Ser Ser Ala Leu Ser Ser Pro Arg Arg Ser Leu Gly Leu	
gca ctc acc cat tcc ttc ggc ccc agt ctt gca gac aca gac ctg tca	4512
Ala Leu Thr His Ser Phe Gly Pro Ser Leu Ala Asp Thr Asp Leu Ser	
ccc atg gat ggc atc agt act tgt ggt cca aag gag gaa gtg acc ctc	4560
Pro Met Asp Gly Ile Ser Thr Cys Gly Pro Lys Glu Glu Val Thr Leu	
cgg gtg gtg gtg agg atg ccc ccg cag cac atc atc aaa ggg gac ttg	4608
Arg Val Val Arg Met Pro Pro Gln His Ile Ile Lys Gly Asp Leu	

aag cag cag gaa ttc ttc ctg ggc tgt agc aag gtc agt gga aaa gtt	4656
Lys Gln Gln Glu Phe Phe Leu Gly Cys Ser Lys Val Ser Gly Lys Val	
gac tgg aag atg ctg gat gaa gct gtt ttc caa gtg ttc aag gac tat	4704
Asp Trp Lys Met Leu Asp Glu Ala Val Phe Gln Val Phe Lys Asp Tyr	
att tct aaa atg gac cca gcc tct acc ctg gga cta agc act gag tcc	4752
Ile Ser Lys Met Asp Pro Ala Ser Thr Leu Gly Leu Ser Thr Glu Ser	
atc cat ggc tac agc atc agc cac gtg aaa cga gtg ttg gat gca gag	4800
Ile His Gly Tyr Ser Ile Ser His Val Lys Arg Val Leu Asp Ala Glu	
ccc ccc gag atg cct cct tgc cgt cga ggt gtc aat aac ata tca gtc	4848
Pro Pro Glu Met Pro Pro Cys Arg Arg Gly Val Asn Asn Ile Ser Val	
tcc ctc aaa ggt ctg aag gag aaa tgc gtc gac agc ctg gtg ttc gag	4896
Ser Leu Lys Gly Leu Lys Glu Lys Cys Val Asp Ser Leu Val Phe Glu	
acg ctg atc ccc aag ccg atg atg cag cac tac ata agc ctc ctg ctg	4944
Thr Leu Ile Pro Lys Pro Met Met Gln His Tyr Ile Ser Leu Leu Leu	
aag cac cgg cgc ctc gtc ctc tcg ggc ccc agc ggc acg ggc aag acc	4992
Lys His Arg Arg Leu Val Leu Ser Gly Pro Ser Gly Thr Gly Lys Thr	
tac ctg acc aat cgc ttg gcc gag tac ctg gtg gag cgc tct ggc cgt	5040
Tyr Leu Thr Asn Arg Leu Ala Glu Tyr Leu Val Glu Arg Ser Gly Arg	
gag gtc aca gag ggc atc gtc agc acc ttc aac atg cac cag cag tct	5088
Glu Val Thr Glu Gly Ile Val Ser Thr Phe Asn Met His Gln Gln Ser	
tgc aag gat ctg caa ctg tat ctt tcc aac cta gcc aac cag ata gac	5136
Cys Lys Asp Leu Gln Leu Tyr Leu Ser Asn Leu Ala Asn Gln Ile Asp	
cgg gaa aca gga att ggg gat gtg ccc ctg gtg att cta ttg gat gac	5184
Arg Glu Thr Gly Ile Gly Asp Val Pro Leu Val Ile Leu Leu Asp Asp	
ctg agt gaa gca ggc tcc atc agt gag ttg gtc aat ggg gcc ctc acc	5232
Leu Ser Glu Ala Gly Ser Ile Ser Glu Leu Val Asn Gly Ala Leu Thr	
tgc aag tat cat aaa tgt ccc tat att ata ggt acc acc aat cag cct	5280
Cys Lys Tyr His Lys Cys Pro Tyr Ile Ile Gly Thr Thr Asn Gln Pro	
gta aaa atg aca ccc aac cat ggc ttt cac ttg agc ttc agg atg ttg	5328
Val Lys Met Thr Pro Asn His Gly Phe His Leu Ser Phe Arg Met Leu	
acc ttc tcc aac aac gtg gag cca gcc aat ggc ttc ctg gtt cgt tac	5376
Thr Phe Ser Asn Asn Val Glu Pro Ala Asn Gly Phe Leu Val Arg Tyr	
ctg agg agg aag ctg gta gag tca gac agc gac atc aat gcc aac aag	5424
Leu Arg Arg Lys Leu Val Glu Ser Asp Ser Asp Ile Asn Ala Asn Lys	
gaa gag ctg ctt cgg gtg ctc gac tgg gta ccc aag ctg tgg tat cat	5472
Glu Glu Leu Leu Arg Val Leu Asp Trp Val Pro Lys Leu Trp Tyr His	
ctc cac acc ttc ctt gag aag cac agc acc tca gac ttc ctc atc ggc	5520

Leu His Thr Phe Leu Glu Lys His Ser Thr Ser Asp Phe Leu Ile Gly	
cct tgc ttc ttt ctg tcg tgt ccc att ggc att gag gac ttc cg acc	5568
Pro Cys Phe Phe Leu Ser Cys Pro Ile Gly Ile Glu Asp Phe Arg Thr	
tgg ttc att gac ctg tgg aac aac tct atc att ccc tat cta cag gaa	5616
Trp Phe Ile Asp Leu Trp Asn Asn Ser Ile Ile Pro Tyr Leu Gln Glu	
gga gcc aag gat ggg ata aag gtc cat gga cag aaa gct gct tgg gag	5664
Gly Ala Lys Asp Gly Ile Lys Val His Gly Gln Lys Ala Ala Trp Glu	
gac cca gtg gaa tgg gtc cg gac aca ctt ccc tgg cca tca gcc caa	5712
Asp Pro Val Glu Trp Val Arg Asp Thr Leu Pro Trp Pro Ser Ala Gln	
caa gac caa tca aag ctg tac cac ctg ccc cca ccc acc gtg ggc cct	5760
Gln Asp Gln Ser Lys Leu Tyr His Leu Pro Pro Thr Val Gly Pro	
cac agc att gcc tca cct ccc gag gat agg aca gtc aaa gac agc acc	5808
His Ser Ile Ala Ser Pro Pro Glu Asp Arg Thr Val Lys Asp Ser Thr	
cca agt tct ctg gac tca gat cct ctg atg gcc atg ctg ctg aaa ctt	5856
Pro Ser Ser Leu Asp Ser Pro Leu Met Ala Met Leu Leu Lys Leu	
caa gaa gct gcc aac tac att gag tct cca gat cga gaa acc atc ctg	5904
Gln Glu Ala Ala Asn Tyr Ile Glu Ser Pro Asp Arg Glu Thr Ile Leu	
gac ccc aac ctt cag gca aca ctt taagggtcg gcaatcactg	5948
Asp Pro Asn Leu Gln Ala Thr Leu	
tcaccccccgg acagcagaac gctggcatca gctatcttag ctcctcctct cccctctcct	6008
ctttcagagc actggctctc cagccccagg aggagaacag gagggaggag gagatgaaag	6068
aggagggaca ggttcttggt gctgtacctt tgagaacttc ctaggaagga atggtggggt	6128
ggcgtttggg aacttgcgcc ccctaaacac atttactggc ctcctctaatt gactttgggg	6188
aaaagatgtat tctgggtctt tcccttgact tcttgggtca attacaaact cctgggcttt	6248
ctggggaggg gttcagaaaa cataaaaaca ctgcagcagt tcctaaatga ttctcacaag	6308
caacccttag agagacagtc ttgtgagggg gatctggggg aggcaggaag ctcctcagat	6368
tttctcacag acccttccca attccatcac cactgccaac aactcctccc ccagagatct	6428
ggctggagcc cagaaaaaga agcatgtggt ttaaaaatg tttaaatcaa tctgtaaaag	6488
gtaaaaatga aaaacaaaaa caagcaaaca aacaaaaaac aatggaaaag atgaagctgg	6548
agagagagga accagttgcc aaggtagaga gctgcccgt cctgcccctct ggatgacata	6608
ggggacatca acaagacggc tgccaacctg agaagtcacc aaaccacaaa aataacctt	6668
cagccttcag ggaaagacta ccagctctgt ctttctaccc tctaattaa caatgcataa	6728
gagtcaataa accctacttt ttaaaaaaaaaaaaaaaag	6768

<210> 24  
<211> 11  
<212> PRT  
<213> Homo sapiens

<400> 24

Arg Gly Ala Ser Ile Leu Pro Leu Thr Asp Phe  
1 5 10

<210> 25  
<211> 1976  
<212> PRT  
<213> Homo sapiens

<400> 25

Ile Leu Ile Cys Ile Phe Pro Ala Ala Pro Pro Leu Phe Leu Arg Pro  
5 10 15

Arg Pro Ile Ala Pro Arg Leu Pro Cys Ser Phe Leu Phe Pro Gly Phe  
20 25 30

Leu Pro Arg Val Ser Phe Pro Cys Ala Leu Gly Leu Pro Leu Ser Leu  
35 40 45

Leu Pro Arg Ser Leu Pro Leu Leu Ser Pro Ser Ser Val Ser Ser  
50 55 60

Val Leu Ser Leu Pro Leu Leu Leu Pro Arg Leu Leu Leu Arg Ser  
65 70 75 80

Arg Pro Leu Pro Pro Pro Val Pro Ala Asp Ala Arg Ile Val His  
85 90 95

Ala Leu Leu Ala Gly Arg Met Leu Gly Ser Ser Val Lys Ser Val Gln  
100 105 110

Pro Glu Val Glu Leu Ser Ser Gly Gly Asp Glu Gly Ala Asp Glu  
115 120 125

Pro Arg Gly Ala Gly Arg Lys Ala Ala Ala Asp Gly Arg Gly Met  
130 135 140

Leu Pro Lys Arg Ala Lys Ala Pro Gly Gly Gly Gly Met Ala Lys  
145 150 155 160

Ala Ser Ala Ala Glu Leu Lys Val Phe Lys Ser Gly Ser Val Asp Ser

165	170	175
Arg Val Pro Gly Gly Pro Pro Ala Ser Asn Leu Arg Lys Gln Lys Ser		
180	185	190
Leu Thr Asn Leu Ser Phe Leu Thr Asp Ser Glu Lys Lys Leu Gln Leu		
195	200	205
Tyr Glu Pro Glu Trp Ser Asp Asp Met Ala Lys Ala Pro Lys Gly Leu		
210	215	220
Gly Lys Val Gly Ser Lys Gly Arg Glu Ala Pro Leu Met Ser Lys Thr		
225	230	235
Leu Ser Lys Ser Glu His Ser Leu Phe Gln Ala Lys Gly Ser Pro Ala		
245	250	255
Gly Gly Ala Lys Thr Pro Leu Ala Pro Leu Ala Pro Asn Leu Gly Lys		
260	265	270
Pro Ser Arg Ile Pro Arg Gly Pro Tyr Ala Glu Val Lys Pro Leu Ser		
275	280	285
Lys Ala Pro Glu Ala Ala Val Ser Glu Asp Gly Lys Ser Asp Asp Glu		
290	295	300
Leu Leu Ser Ser Lys Ala Lys Ala Gln Lys Ser Ser Gly Pro Val Pro		
305	310	315
Ser Ala Lys Gly Gln Glu Glu Arg Ala Phe Leu Lys Val Asp Pro Glu		
325	330	335
Leu Val Val Thr Val Leu Gly Asp Leu Glu Gln Leu Leu Phe Ser Gln		
340	345	350
Met Leu Asp Pro Glu Ser Gln Arg Lys Arg Thr Val Gln Asn Val Leu		
355	360	365
Asp Leu Arg Gln Asn Leu Glu Glu Thr Met Ser Ser Leu Arg Gly Ser		
370	375	380
Gln Val Thr His Ser Ser Leu Glu Met Thr Cys Tyr Asp Ser Asp Asp		
385	390	395
400		
Ala Asn Pro Arg Ser Val Ser Ser Leu Ser Asn Arg Ser Tyr Pro Leu		
405	410	415
Ser Trp Arg Tyr Gly Gln Ser Ser Pro Arg Leu Gln Ala Gly Asp Ala		
420	425	430
Pro Ser Val Gly Gly Ser Cys Arg Ser Glu Gly Thr Pro Ala Trp Tyr		
435	440	445
Met His Gly Glu Arg Ala His Tyr Ser His Thr Met Pro Met Arg Ser		
450	455	460

Pro	Ser	Lys	Leu	Ser	His	Ile	Ser	Arg	Leu	Glu	Leu	Val	Glu	Ser	Leu
465															480
Asp	Ser	Asp	Glu	Val	Asp	Leu	Lys	Ser	Gly	Tyr	Met	Ser	Asp	Ser	Asp
															495
Leu	Met	Gly	Lys	Thr	Met	Thr	Glu	Asp	Asp	Asp	Ile	Thr	Thr	Gly	Trp
															510
Asp	Glu	Ser	Ser	Ser	Ile	Ser	Ser	Gly	Leu	Ser	Asp	Ala	Ser	Asp	Asn
															525
Leu	Ser	Ser	Glu	Glu	Phe	Asn	Ala	Ser	Ser	Ser	Leu	Asn	Ser	Leu	Pro
															540
Ser	Thr	Pro	Thr	Ala	Ser	Arg	Arg	Asn	Ser	Thr	Ile	Val	Leu	Arg	Thr
															560
Asp	Ser	Glu	Lys	Arg	Ser	Leu	Ala	Glu	Ser	Gly	Leu	Ser	Trp	Phe	Ser
															575
Glu	Ser	Glu	Glu	Lys	Ala	Pro	Lys	Lys	Leu	Glu	Tyr	Asp	Ser	Gly	Ser
															590
Leu	Lys	Met	Glu	Pro	Gly	Thr	Ser	Lys	Trp	Arg	Arg	Glu	Arg	Pro	Glu
															605
Ser	Cys	Asp	Asp	Ser	Ser	Lys	Gly	Gly	Glu	Leu	Lys	Lys	Pro	Ile	Ser
															620
Leu	Gly	His	Pro	Gly	Ser	Leu	Lys	Lys	Gly	Lys	Thr	Pro	Pro	Val	Ala
															640
Val	Thr	Ser	Pro	Ile	Thr	His	Thr	Ala	Gln	Ser	Ala	Leu	Lys	Val	Ala
															655
Gly	Lys	Pro	Glu	Gly	Lys	Ala	Thr	Asp	Lys	Gly	Lys	Leu	Ala	Val	Lys
															670
Asn	Thr	Gly	Leu	Gln	Arg	Ser	Ser	Asp	Ala	Gly	Arg	Asp	Arg	Leu	
															685
Ser	Asp	Ala	Lys	Lys	Pro	Pro	Ser	Gly	Ile	Ala	Arg	Pro	Ser	Thr	Ser
															700
Gly	Ser	Phe	Gly	Tyr	Lys	Lys	Pro	Pro	Pro	Ala	Thr	Gly	Thr	Ala	Thr
															720
Val	Met	Gln	Thr	Gly	Gly	Ser	Ala	Thr	Leu	Ser	Lys	Ile	Gln	Lys	Ser
															735
Ser	Gly	Ile	Pro	Val	Lys	Pro	Val	Asn	Gly	Arg	Lys	Thr	Ser	Leu	Asp
															750
Val	Ser	Asn	Ser	Ala	Glu	Pro	Gly	Phe	Leu	Ala	Pro	Gly	Ala	Arg	Ser
															765

Asn	Ile	Gln	Tyr	Arg	Ser	Leu	Pro	Arg	Pro	Ala	Lys	Ser	Ser	Ser	Met
770						775				780					
Ser	Val	Thr	Gly	Gly	Arg	Gly	Gly	Pro	Arg	Pro	Val	Ser	Ser	Ser	Ile
785						790				795					800
Asp	Pro	Ser	Leu	Leu	Ser	Thr	Lys	Gln	Gly	Gly	Leu	Thr	Pro	Ser	Arg
						805				810					815
Leu	Lys	Glu	Pro	Thr	Lys	Val	Ala	Ser	Gly	Arg	Thr	Thr	Pro	Ala	Pro
						820				825					830
Val	Asn	Gln	Thr	Asp	Arg	Glu	Lys	Glu	Lys	Ala	Lys	Ala	Lys	Ala	Val
						835				840					845
Ala	Leu	Asp	Ser	Asp	Asn	Ile	Ser	Leu	Lys	Ser	Ile	Gly	Ser	Pro	Glu
						850				855					860
Ser	Thr	Pro	Lys	Asn	Gln	Ala	Ser	His	Pro	Thr	Ala	Thr	Lys	Leu	Ala
						865				870					880
Glu	Leu	Pro	Pro	Thr	Pro	Leu	Arg	Ala	Thr	Ala	Lys	Ser	Phe	Val	Lys
						885				890					895
Pro	Pro	Ser	Leu	Ala	Asn	Leu	Asp	Lys	Val	Asn	Ser	Asn	Ser	Leu	Asp
						900				905					910
Leu	Pro	Ser	Ser	Ser	Asp	Thr	Thr	His	Ala	Ser	Lys	Val	Pro	Asp	Leu
						915				920					925
His	Ala	Thr	Ser	Ser	Ala	Ser	Gly	Gly	Pro	Leu	Pro	Ser	Cys	Phe	Thr
						930				935					940
Pro	Ser	Pro	Ala	Pro	Ile	Leu	Asn	Ile	Asn	Ser	Ala	Ser	Phe	Ser	Gln
						945				950					960
Gly	Leu	Glu	Leu	Met	Ser	Gly	Phe	Ser	Val	Pro	Lys	Glu	Thr	Arg	Met
						965				970					975
Tyr	Pro	Lys	Leu	Ser	Gly	Leu	His	Arg	Ser	Met	Glu	Ser	Leu	Gln	Met
						980				985					990
Pro	Met	Ser	Leu	Pro	Ser	Ala	Phe	Pro	Ser	Ser	Thr	Pro	Val	Pro	Thr
						995				1000					1005
Pro	Pro	Ala	Pro	Pro	Ala	Ala	Pro	Thr	Glu	Glu	Glu	Thr	Glu	Glu	Leu
						1010				1015					1020
Thr	Trp	Ser	Gly	Ser	Pro	Arg	Ala	Gly	Gln	Leu	Asp	Ser	Asn	Gln	Arg
						1025				1030					1040
Asp	Arg	Asn	Thr	Leu	Pro	Lys	Lys	Gly	Leu	Arg	Tyr	Gln	Leu	Gln	Ser
						1045				1050					1055
Gln	Glu	Glu	Thr	Lys	Glu	Arg	Arg	His	Ser	His	Thr	Ile	Gly	Gly	Leu
						1060				1065					1070

Pro	Glu	Ser	Asp	Asp	Gln	Ser	Glu	Leu	Pro	Ser	Pro	Pro	Ala	Leu	Pro
1075							1080						1085		
Met	Ser	Leu	Ser	Ala	Lys	Gly	Gln	Leu	Thr	Asn	Ile	Val	Ser	Pro	Thr
1090						1095					1100				
Ala	Ala	Thr	Thr	Pro	Arg	Ile	Thr	Arg	Ser	Asn	Ser	Ile	Pro	Thr	His
1105						1110				1115			1120		
Glu	Ala	Ala	Phe	Glu	Leu	Tyr	Ser	Gly	Ser	Gln	Met	Gly	Ser	Thr	Leu
			1125					1130					1135		
Ser	Leu	Ala	Glu	Arg	Pro	Lys	Gly	Met	Ile	Arg	Ser	Gly	Ser	Phe	Arg
			1140				1145					1150			
Asp	Pro	Thr	Asp	Asp	Val	His	Gly	Ser	Val	Leu	Ser	Leu	Ala	Ser	Ser
			1155			1160					1165				
Ala	Ser	Ser	Thr	Tyr	Ser	Ser	Ala	Glu	Glu	Arg	Met	Gln	Ser	Glu	Gln
			1170			1175				1180					
Ile	Arg	Lys	Leu	Arg	Arg	Glu	Leu	Glu	Ser	Ser	Gln	Glu	Lys	Val	Ala
			1185			1190				1195			1200		
Thr	Leu	Thr	Ser	Gln	Leu	Ser	Ala	Asn	Ala	Asn	Leu	Val	Ala	Ala	Phe
			1205				1210					1215			
Glu	Gln	Ser	Leu	Val	Asn	Met	Thr	Ser	Arg	Leu	Arg	His	Leu	Ala	Glu
			1220			1225					1230				
Thr	Ala	Glu	Glu	Lys	Asp	Thr	Glu	Leu	Leu	Asp	Leu	Arg	Glu	Thr	Ile
			1235				1240				1245				
Asp	Phe	Leu	Lys	Lys	Asn	Ser	Glu	Ala	Gln	Ala	Val	Ile	Gln	Gly	
			1250			1255				1260					
Ala	Leu	Asn	Ala	Ser	Glu	Thr	Thr	Pro	Lys	Glu	Leu	Arg	Ile	Lys	Arg
			1265			1270				1275			1280		
Gln	Asn	Ser	Ser	Asp	Ser	Ile	Ser	Ser	Leu	Asn	Ser	Ile	Thr	Ser	His
			1285				1290					1295			
Ser	Ser	Ile	Gly	Ser	Ser	Lys	Asp	Ala	Asp	Ala	Lys	Lys	Lys	Lys	
			1300			1305					1310				
Lys	Ser	Trp	Leu	Arg	Ser	Ser	Phe	Asn	Lys	Ala	Phe	Ser	Ile	Lys	Lys
			1315				1320				1325				
Gly	Pro	Lys	Ser	Ala	Ser	Ser	Tyr	Ser	Asp	Ile	Glu	Glu	Ile	Ala	Thr
			1330				1335				1340				
Pro	Asp	Ser	Ser	Ala	Pro	Ser	Ser	Pro	Lys	Leu	Gln	His	Gly	Ser	Thr
			1345				1350				1355			1360	
Glu	Thr	Ala	Ser	Pro	Ser	Ile	Lys	Ser	Ser	Thr	Ser	Ser	Val	Gly	

1365	1370	1375
Thr Asp Val Thr Glu Gly Pro Ala His Pro Ala Pro His Thr Arg Leu		
1380	1385	1390
Phe His Ala Asn Glu Glu Glu Pro Glu Lys Lys Glu Val Ser Glu		
1395	1400	1405
Leu Arg Ser Glu Leu Trp Glu Lys Glu Met Lys Leu Thr Asp Ile Arg		
1410	1415	1420
Leu Glu Ala Leu Asn Ser Ala His Gln Leu Asp Gln Leu Arg Glu Thr		
1425	1430	1435
Met His Asn Met Gln Leu Glu Val Asp Leu Leu Glu Ala Glu Asn Asp		
1445	1450	1455
Arg Leu Lys Val Ala Pro Gly Pro Ser Ser Gly Ser Thr Pro Gly Gln		
1460	1465	1470
Val Pro Gly Ser Ser Ala Leu Ser Ser Pro Arg Arg Ser Leu Gly Leu		
1475	1480	1485
Ala Leu Thr His Ser Phe Gly Pro Ser Leu Ala Asp Thr Asp Leu Ser		
1490	1495	1500
Pro Met Asp Gly Ile Ser Thr Cys Gly Pro Lys Glu Glu Val Thr Leu		
1505	1510	1515
1520		
Arg Val Val Val Arg Met Pro Pro Gln His Ile Ile Lys Gly Asp Leu		
1525	1530	1535
Lys Gln Gln Glu Phe Phe Leu Gly Cys Ser Lys Val Ser Gly Lys Val		
1540	1545	1550
Asp Trp Lys Met Leu Asp Glu Ala Val Phe Gln Val Phe Lys Asp Tyr		
1555	1560	1565
Ile Ser Lys Met Asp Pro Ala Ser Thr Leu Gly Leu Ser Thr Glu Ser		
1570	1575	1580
Ile His Gly Tyr Ser Ile Ser His Val Lys Arg Val Leu Asp Ala Glu		
1585	1590	1595
1600		
Pro Pro Glu Met Pro Pro Cys Arg Arg Gly Val Asn Asn Ile Ser Val		
1605	1610	1615
Ser Leu Lys Gly Leu Lys Glu Lys Cys Val Asp Ser Leu Val Phe Glu		
1620	1625	1630
Thr Leu Ile Pro Lys Pro Met Met Gln His Tyr Ile Ser Leu Leu Leu		
1635	1640	1645
Lys His Arg Arg Leu Val Leu Ser Gly Pro Ser Gly Thr Gly Lys Thr		
1650	1655	1660

Tyr Leu Thr Asn Arg Leu Ala Glu Tyr Leu Val Glu Arg Ser Gly Arg  
 1665 1670 1675 1680

Glu Val Thr Glu Gly Ile Val Ser Thr Phe Asn Met His Gln Gln Ser  
 1685 1690 1695

Cys Lys Asp Leu Gln Leu Tyr Leu Ser Asn Leu Ala Asn Gln Ile Asp  
 1700 1705 1710

Arg Glu Thr Gly Ile Gly Asp Val Pro Leu Val Ile Leu Leu Asp Asp  
 1715 1720 1725

Leu Ser Glu Ala Gly Ser Ile Ser Glu Leu Val Asn Gly Ala Leu Thr  
 1730 1735 1740

Cys Lys Tyr His Lys Cys Pro Tyr Ile Ile Gly Thr Thr Asn Gln Pro  
 1745 1750 1755 1760

Val Lys Met Thr Pro Asn His Gly Phe His Leu Ser Phe Arg Met Leu  
 1765 1770 1775

Thr Phe Ser Asn Asn Val Glu Pro Ala Asn Gly Phe Leu Val Arg Tyr  
 1780 1785 1790

Leu Arg Arg Lys Leu Val Glu Ser Asp Ser Asp Ile Asn Ala Asn Lys  
 1795 1800 1805

Glu Glu Leu Leu Arg Val Leu Asp Trp Val Pro Lys Leu Trp Tyr His  
 1810 1815 1820

Leu His Thr Phe Leu Glu Lys His Ser Thr Ser Asp Phe Leu Ile Gly  
 1825 1830 1835 1840

Pro Cys Phe Phe Leu Ser Cys Pro Ile Gly Ile Glu Asp Phe Arg Thr  
 1845 1850 1855

Trp Phe Ile Asp Leu Trp Asn Asn Ser Ile Ile Pro Tyr Leu Gln Glu  
 1860 1865 1870

Gly Ala Lys Asp Gly Ile Lys Val His Gly Gln Lys Ala Ala Trp Glu  
 1875 1880 1885

Asp Pro Val Glu Trp Val Arg Asp Thr Leu Pro Trp Pro Ser Ala Gln  
 1890 1895 1900

Gln Asp Gln Ser Lys Leu Tyr His Leu Pro Pro Pro Thr Val Gly Pro  
 1905 1910 1915 1920

His Ser Ile Ala Ser Pro Pro Glu Asp Arg Thr Val Lys Asp Ser Thr  
 1925 1930 1935

Pro Ser Ser Leu Asp Ser Asp Pro Leu Met Ala Met Leu Leu Lys Leu  
 1940 1945 1950

Gln Glu Ala Ala Asn Tyr Ile Glu Ser Pro Asp Arg Glu Thr Ile Leu  
 1955 1960 1965

Asp Pro Asn Leu Gln Ala Thr Leu  
1970 1975

<210> 26  
<211> 7783  
<212> DNA  
<213> Homo sapiens

<220>  
<221> CDS  
<223> (2)..(7141)

<400> 26

c aac cag cca gaa cgc ctg aac tcg cag gtg ctg cag ggg ctg cag 46  
Asn Gln Pro Glu Arg Leu Asn Ser Gln Val Leu Gln Gly Leu Gln

gag cca gcg ggg gag ggg ctc ccg ctg cgg aag agc ggc tcg gtg gaa 94  
Glu Pro Ala Gly Glu Gly Leu Pro Leu Arg Lys Ser Gly Ser Val Glu

aac ggg ttc gat acc cag atc tac aca gac tgg gcc aat cat tac cta 142  
Asn Gly Phe Asp Thr Gln Ile Tyr Thr Asp Trp Ala Asn His Tyr Leu

gcc aaa tcc ggc cac aag cgt ctc atc agg gat ctc cag caa gat gtg 190  
Ala Lys Ser Gly His Lys Arg Leu Ile Arg Asp Leu Gln Gln Asp Val

aca gat ggc gtc ctc ctg gcc cag att atc cag gtt gtg gca aat gaa 238  
Thr Asp Gly Val Leu Leu Ala Gln Ile Ile Gln Val Val Ala Asn Glu

aag att gaa gac atc aat ggc tgt ccg aag aac aga tcc caa atg att 286  
Lys Ile Glu Asp Ile Asn Gly Cys Pro Lys Asn Arg Ser Gln Met Ile

gaa aac ata gat gcc tgc ttg aat ttc ctg gca gct aag gga ata aac 334  
Glu Asn Ile Asp Ala Cys Leu Asn Phe Leu Ala Lys Gly Ile Asn

atc cag ggg ctg tct gca gaa gag atc agg aat gga aac ctc aag gcc 382  
Ile Gln Gly Leu Ser Ala Glu Glu Ile Arg Asn Gly Asn Leu Lys Ala

att cta ggc ctc ttc agc ctc tcc cga tac aag cag cag cag cag 430  
Ile Leu Gly Leu Phe Phe Ser Leu Ser Arg Tyr Lys Gln Gln Gln

cag ccc cag aag cag cac ctc tcc tca cct ctg ccg ccc gcc gta tcc 478  
Gln Pro Gln Lys Gln His Leu Ser Ser Pro Leu Pro Pro Ala Val Ser

cag gtg gcc ggg gcc ccc tcc cag tgc cag gct ggc acc cct cag cag	526
Gln Val Ala Gly Ala Pro Ser Gln Cys Gln Ala Gly Thr Pro Gln Gln	
cag gtg cca gtc act ccc caa gcc ccg tgc cag cct cac cag cca gcg	574
Gln Val Pro Val Thr Pro Gln Ala Pro Cys Gln Pro His Gln Pro Ala	
cca cat cag cag tca aaa gca caa gct gaa atg cag tcc aga ctt cca	622
Pro His Gln Gln Ser Lys Ala Gln Ala Glu Met Gln Ser Arg Leu Pro	
ggt cct acc gcg agg gta tcc gct gca ggc agc gag gcc aaa aca cgc	670
Gly Pro Thr Ala Arg Val Ser Ala Ala Gly Ser Glu Ala Lys Thr Arg	
gga ggg tca act act gct aac aac cga cgc agc cag agc ttt aac aac	718
Gly Gly Ser Thr Thr Ala Asn Asn Arg Arg Ser Gln Ser Phe Asn Asn	
tat gat aaa tcc aaa cca gtc acc tcc cca ccc cca ccg cca agc agc	766
Tyr Asp Lys Ser Lys Pro Val Thr Ser Pro Pro Pro Pro Ser Ser	
cac gag aaa gag cct ttg gca agt tca gcc tcc tcc cac ccc gga atg	814
His Glu Lys Glu Pro Leu Ala Ser Ser Ala Ser Ser His Pro Gly Met	
agt gac aat gca cct gct tcc ttg gag agc ggc agc agc tcc acc cct	862
Ser Asp Asn Ala Pro Ala Ser Leu Glu Ser Gly Ser Ser Thr Pro	
act aat tgc agt acc tcc tcg gcc atc ccg cag ccc ggt gca gcc acc	910
Thr Asn Cys Ser Thr Ser Ala Ile Pro Gln Pro Gly Ala Ala Thr	
aag cct tgg cgc agc aaa tcc ctc agc gtg aag cac agt gcc acg gta	958
Lys Pro Trp Arg Ser Lys Ser Leu Ser Val Lys His Ser Ala Thr Val	
tcc atg ctc tcg gtc aag cct cct ggg cct gag gcc ccc agg ccc aca	1006
Ser Met Leu Ser Val Lys Pro Pro Gly Pro Glu Ala Pro Arg Pro Thr	
cct gaa gcc atg aag ccg gcc ccc aac aat cag aag tcc atg ctg gaa	1054
Pro Glu Ala Met Lys Pro Ala Pro Asn Asn Gln Lys Ser Met Leu Glu	
aag ctg aaa ctt ttc aac agt aaa ggg ggc tca aag gca ggt gag ggg	1102
Lys Leu Lys Leu Phe Asn Ser Lys Gly Gly Ser Lys Ala Gly Glu Gly	
ccg ggg tcc cgg gac aca agc tgt gag cgg ctg gag act ctg ccc agc	1150
Pro Gly Ser Arg Asp Thr Ser Cys Glu Arg Leu Glu Thr Leu Pro Ser	

ttc gaa gag agc gag gag ctg gag gcc gcc agt cgc atg ctc acc acc Phe Glu Glu Ser Glu Glu Leu Ala Ala Ser Arg Met Leu Thr Thr	1198
gtg ggc cct gct tcc agc agc ccc aag att gca ctc aag ggc att gcc Val Gly Pro Ala Ser Ser Pro Lys Ile Ala Leu Lys Gly Ile Ala	1246
cag agg act ttt agc cgg gca ctg acc aac aag aag agt tct ctg aaa Gln Arg Thr Phe Ser Arg Ala Leu Thr Asn Lys Lys Ser Ser Leu Lys	1294
ggc aat gag aaa gag aag gag aaa caa cag cgg gag aag gat aag gag Gly Asn Glu Lys Glu Lys Glu Gln Arg Glu Lys Asp Lys Glu	1342
aaa agc aag gac ctt gcc aag aga gcc tct gtg acg gag agg ctg gac Lys Ser Lys Asp Leu Ala Lys Arg Ala Ser Val Thr Glu Arg Leu Asp	1390
ctc aag gag gag cca aaa gaa gac ccc agt gga gca gct gtg ccc gag Leu Lys Glu Glu Pro Lys Glu Asp Pro Ser Gly Ala Ala Val Pro Glu	1438
atg cca aaa aag tcc tcc aag att gcc agc ttc atc ccc aaa ggg ggg Met Pro Lys Lys Ser Ser Lys Ile Ala Ser Phe Ile Pro Lys Gly Gly	1486
aag ctc aac agt gcc aag aag gag ccc atg gcc cct tcc cac agt gga Lys Leu Asn Ser Ala Lys Glu Pro Met Ala Pro Ser His Ser Gly	1534
ata cca aaa cca gga atg aag agc atg ccc ggg aaa tcc cca agt gcc Ile Pro Lys Pro Gly Met Lys Ser Met Pro Gly Lys Ser Pro Ser Ala	1582
cca gcg cct tcc aag gaa ggg gag cgg agc cgg agt ggg aag ctg agc Pro Ala Pro Ser Lys Glu Gly Glu Arg Ser Arg Gly Lys Leu Ser	1630
tca gga ctc ccc cag cag aag ccc cag ctg gac ggc aga cac tcc agt Ser Gly Leu Pro Gln Gln Lys Pro Gln Leu Asp Gly Arg His Ser Ser	1678
tcc tct tcc agc ctg gcg tcc tca gaa gga aaa ggc cca gga ggg acc Ser Ser Ser Leu Ala Ser Ser Glu Gly Lys Pro Gly Gly Thr	1726
acc ctg aac cac agc atc agc agc cag act gtc agt ggg tct gtc ggg Thr Leu Asn His Ser Ile Ser Ser Gln Thr Val Ser Gly Ser Val Gly	1774
acc acc cag acc aca gga agc aat acc gtc agt gtt cag cta cct cag Thr Thr Gln Thr Thr Gly Ser Asn Thr Val Ser Val Gln Leu Pro Gln	1822

ccc cag cag caa tac aac cat ccc aac act gcc acg gtt gca cct ttc	1870
Pro Gln Gln Gln Tyr Asn His Pro Asn Thr Ala Thr Val Ala Pro Phe	
ctg tac agg tct cag acg gac act gaa ggg aat gtt act gcc gag tca	1918
Leu Tyr Arg Ser Gln Thr Asp Thr Glu Gly Asn Val Thr Ala Glu Ser	
agc tca aca ggt gtg agc gtg gag ccc agc cac ttc acc aag act gga	1966
Ser Ser Thr Gly Val Ser Val Glu Pro Ser His Phe Thr Lys Thr Gly	
cag cct gct ctg gaa gaa ctc act ggg gaa gat cct gag gct cgg cgg	2014
Gln Pro Ala Leu Glu Glu Leu Thr Gly Glu Asp Pro Glu Ala Arg Arg	
ctg cgg aca gtg aag aac atc gct gat ctg cgg cag aat ttg gag gaa	2062
Leu Arg Thr Val Lys Asn Ile Ala Asp Leu Arg Gln Asn Leu Glu Glu	
acc atg tcc agt tta agg gga act cag gtt aca cac agc aca ttg gaa	2110
Thr Met Ser Ser Leu Arg Gly Thr Gln Val Thr His Ser Thr Leu Glu	
acc acg ttt gac acc aat gtc acc acg gag atg agt ggc cgt agc ata	2158
Thr Thr Phe Asp Thr Asn Val Thr Thr Glu Met Ser Gly Arg Ser Ile	
ctc agc ttg aca ggg agg ccc aca cct ctg tcc tgg aga ctg ggc cag	2206
Leu Ser Leu Thr Gly Arg Pro Thr Pro Leu Ser Trp Arg Leu Gly Gln	
tcc agc cct cgg ctc caa gca gga gac gcc ccc tca atg ggc aat ggg	2254
Ser Ser Pro Arg Leu Gln Ala Gly Asp Ala Pro Ser Met Gly Asn Gly	
tat ccc cct cga gcc aac gcc agc agg ttc atc aac act gag tca ggt	2302
Tyr Pro Pro Arg Ala Asn Ala Ser Arg Phe Ile Asn Thr Glu Ser Gly	
cgc tat gtg tac tcc gcc cct ctg aga agg cag ctg gcc tcc cgg ggc	2350
Arg Tyr Val Tyr Ser Ala Pro Leu Arg Arg Gln Leu Ala Ser Arg Gly	
agt agt gtc tgc cac gtg gac gtc tca gac aag gca gga gat gag atg	2398
Ser Ser Val Cys His Val Asp Val Ser Asp Lys Ala Gly Asp Glu Met	
gac ctg gaa ggc atc agc atg gac gcc ccc ggc tac atg agc gat ggg	2446
Asp Leu Glu Gly Ile Ser Met Asp Ala Pro Gly Tyr Met Ser Asp Gly	
gat gtt ctg agc aag aac atc cgg acc gat gac att aca agc gga tac	2494
Asp Val Leu Ser Lys Asn Ile Arg Thr Asp Asp Ile Thr Ser Gly Tyr	
atg act gat ggt gga ctt ggc ctc tat acc cgt cgc ctg aac cgg ctc	2542

Met Thr Asp Gly Gly Leu Gly Leu Tyr Thr Arg Arg Leu Asn Arg Leu	
cct gat ggg atg gct gtg gta cgg gag acc ctg caa cga aat acc tcc Pro Asp Gly Met Ala Val Val Arg Glu Thr Leu Gln Arg Asn Thr Ser	2590
ctg ggc ctc gga gac gct gac agc tgg gac gac agc agc tcc gtc agc Leu Gly Leu Gly Asp Ala Asp Ser Trp Asp Asp Ser Ser Val Ser	2638
agc ggc atc agc gac acc ata gac aac ctc agc act gat gac atc aac Ser Gly Ile Ser Asp Thr Ile Asp Asn Leu Ser Thr Asp Asp Ile Asn	2686
acc agc tcc tcc atc agc tct tat gcc aac aca cct gcc tcc tct cga Thr Ser Ser Ser Ile Ser Ser Tyr Ala Asn Thr Pro Ala Ser Ser Arg	2734
aaa aac ctg gat gtg cag act gat gct gag aag cac tca cag gtg gag Lys Asn Leu Asp Val Gln Thr Asp Ala Glu Lys His Ser Gln Val Glu	2782
agg aat tcc ctg tgg tct ggt gat gat gtc aag aaa tca gac gga ggc Arg Asn Ser Leu Trp Ser Gly Asp Asp Val Lys Lys Ser Asp Gly Gly	2830
tca gac agc ggc ata aaa atg gag cca ggt tcc aag tgg agg cgg aat Ser Asp Ser Gly Ile Lys Met Glu Pro Gly Ser Lys Trp Arg Arg Asn	2878
cct tct gat gtg tct gac gag tcc gac aaa agc acg tcg ggc aag aag Pro Ser Asp Val Ser Asp Glu Ser Asp Lys Ser Thr Ser Gly Lys Lys	2926
aat cct gtc atc tcc cag aca ggc tca tgg cgg cga ggc atg aca gct Asn Pro Val Ile Ser Gln Thr Gly Ser Trp Arg Arg Gly Met Thr Ala	2974
cag gtg ggc atc acc atg cca agg acg aag gct tca gcc ccg gca ggc Gln Val Gly Ile Thr Met Pro Arg Thr Lys Ala Ser Ala Pro Ala Gly	3022
gca ctg aag acc cca gga act gga aaa aca gac gac gca aag gtg tct Ala Leu Lys Thr Pro Gly Thr Lys Thr Asp Asp Ala Lys Val Ser	3070
gag aaa gga agg ctt tct cct aaa gcc tcc cag gtg aag cgc tcc cca Glu Lys Gly Arg Leu Ser Pro Lys Ala Ser Gln Val Lys Arg Ser Pro	3118
tca gat gca ggc cgg agc agt ggt gac gaa tcc aaa aag ccc ctc ccc Ser Asp Ala Gly Arg Ser Ser Gly Asp Glu Ser Lys Lys Pro Leu Pro	3166
agc agc tct agg aca cct act gcc aat gcc aac agc ttt ggg ttc aag	3214

Ser Ser Ser Arg Thr Pro Thr Ala Asn Ala Asn Ser Phe Gly Phe Lys  
aag cag agt ggt tcc gcc acc ggc ctg gcc atg atc aca gcc agc ggg 3262  
Lys Gln Ser Gly Ser Ala Thr Gly Leu Ala Met Ile Thr Ala Ser Gly

gtg act gtc acc agc agg tca gcc aca ctg ggc aaa atc cca aag tca 3310  
Val Thr Val Thr Ser Arg Ser Ala Thr Leu Gly Lys Ile Pro Lys Ser

tct gca ctc gtc agt cgg tct gct ggt cgg aag tca agt atg gat ggg 3358  
Ser Ala Leu Val Ser Arg Ser Ala Gly Arg Lys Ser Ser Met Asp Gly

gct cag aat cag gat gac ggg tat cta gcc cta agc tcc cgg aca aac 3406  
Ala Gln Asn Gln Asp Asp Gly Tyr Leu Ala Leu Ser Ser Arg Thr Asn

ctt cag tac cgg agt ttg ccg agg ccc agt aag tcc aac agc cgg aac 3454  
Leu Gln Tyr Arg Ser Leu Pro Arg Pro Ser Lys Ser Asn Ser Arg Asn

ggg gct ggg aac agg tct agc acc agc agc ata gat tcc aac att agc 3502  
Gly Ala Gly Asn Arg Ser Ser Thr Ser Ser Ile Asp Ser Asn Ile Ser

agc aag tcc gca ggc ctg cca gtg ccc aaa ctg agg gag cct tcc aaa 3550  
Ser Lys Ser Ala Gly Leu Pro Val Pro Lys Leu Arg Glu Pro Ser Lys

aca gcc cta ggc agc tct cta cca ggt ctg gtc aac caa aca gac aag 3598  
Thr Ala Leu Gly Ser Ser Leu Pro Gly Leu Val Asn Gln Thr Asp Lys

gag aaa ggc atc tca tca gac aac gag agt gtg gct tcc tgt aac tcg 3646  
Glu Lys Gly Ile Ser Ser Asp Asn Glu Ser Val Ala Ser Cys Asn Ser

gtg aaa gtg aat ccg gca gcc cag cct gtg tcc agt ccg gct cag acc 3694  
Val Lys Val Asn Pro Ala Ala Gln Pro Val Ser Ser Pro Ala Gln Thr

agt ctc cag cct gga gcc aag tac cca gat gtg gcc tct ccc aca ctc 3742  
Ser Leu Gln Pro Gly Ala Lys Tyr Pro Asp Val Ala Ser Pro Thr Leu

cgc aga ctc ttt ggt ggg aag cct acc aag caa gtg ccc atc gcc aca 3790  
Arg Arg Leu Phe Gly Gly Lys Pro Thr Lys Gln Val Pro Ile Ala Thr

gct gaa aac atg aaa aat tcg gtg gtc atc tcc aat cct cat gcc acc 3838  
Ala Glu Asn Met Lys Asn Ser Val Val Ile Ser Asn Pro His Ala Thr

atg act cag caa ggt aac cta gac tcc ccg tca ggc agt ggc gtc ctg 3886  
Met Thr Gln Gln Gly Asn Leu Asp Ser Pro Ser Gly Ser Gly Val Leu

agc agt ggg agc agt cct ctc tac agc aag aat gtg gac ctc aac Ser Ser Gly Ser Ser Pro Leu Tyr Ser Lys Asn Val Asp Leu Asn	3934
cag tct ccg cta gcc tcc agc ccc agc tca gcc cac tcg gcc cct tcc Gln Ser Pro Leu Ala Ser Pro Ser Ser Ala His Ser Ala Pro Ser	3982
aac agc ctc acc tgg ggc acc aac gcc agc agc tcc tcc gca gtt agc Asn Ser Leu Thr Trp Gly Thr Asn Ala Ser Ser Ser Ala Val Ser	4030
aag gat ggc ctg ggc ttt cag tct gtc agc agc ctc cac acc agc tgt Lys Asp Gly Leu Gly Phe Gln Ser Val Ser Ser Leu His Thr Ser Cys	4078
gag tcc atc gac atc tcc ctc agc agt gga ggg gtc ccc agc cac aat Glu Ser Ile Asp Ile Ser Leu Ser Ser Gly Gly Val Pro Ser His Asn	4126
tct tcc act ggc ctc atc gcc tcc tcc aag gac gac tcc ttg act ccc Ser Ser Thr Gly Leu Ile Ala Ser Ser Lys Asp Asp Ser Leu Thr Pro	4174
ttt gtc aga act aac agt gtg aag acc aca ctg tca gaa agc cct ctc Phe Val Arg Thr Asn Ser Val Lys Thr Thr Leu Ser Glu Ser Pro Leu	4222
tct tcc cct gct agc cct aag ttc tgc aga agt act ctg ccc agg Ser Ser Pro Ala Ala Ser Pro Lys Phe Cys Arg Ser Thr Leu Pro Arg	4270
aaa cag gac agt gac ccg cac ctt gat agg aac act ttg cct aag aaa Lys Gln Asp Ser Asp Pro His Leu Asp Arg Asn Thr Leu Pro Lys Lys	4318
gga ctc agg tat act ccc acc tcc cag ctt cgc acg caa gaa gat gca Gly Leu Arg Tyr Thr Pro Ser Gln Leu Arg Thr Gln Glu Asp Ala	4366
aaa gaa tgg tta cgg tcc cat tct gca gga ggc ctt cag gac acc gct Lys Glu Trp Leu Arg Ser His Ser Ala Gly Leu Gln Asp Thr Ala	4414
gcc aat tcc ccc ttt tcc tct ggc tcc agc gtg act tct ccc tcc gga Ala Asn Ser Pro Phe Ser Ser Gly Ser Val Thr Ser Pro Ser Gly	4462
aca aga ttc aac ttt tcc cag ctt gcg agt ccc acc act gtc acc cag Thr Arg Phe Asn Phe Ser Gln Leu Ala Ser Pro Thr Thr Val Thr Gln	4510
atg agc ttg tcc aac ccg acc atg ctg agg act cac agc ctc tcc aat Met Ser Leu Ser Asn Pro Thr Met Leu Arg Thr His Ser Leu Ser Asn	4558

gct gat ggg cag tat gat cca tac act gac agc cgc ttc cg <sup>g</sup> aat agc Ala Asp Gly Gln Tyr Asp Pro Tyr Thr Asp Ser Arg Phe Arg Asn Ser	4606
tcc atg tcc ctg gat gag aag agc aga acc atg agc cgt tca ggc tca Ser Met Ser Leu Asp Glu Lys Ser Arg Thr Met Ser Arg Ser Gly Ser	4654
ttc cgg gat ggg ttt gaa gaa gtt cat gga tcc tca ctc tcc ttg gtt Phe Arg Asp Gly Phe Glu Val His Gly Ser Ser Leu Ser Leu Val	4702
tcc agc aca tcg tca gtt tat tct aca cca gaa gaa aaa tgc cag tca Ser Ser Thr Ser Val Tyr Ser Thr Pro Glu Glu Lys Cys Gln Ser	4750
gag att cgc aag ctg cgg cgg gaa ctg gat gcc tcc cag gag aaa gtt Glu Ile Arg Lys Leu Arg Arg Glu Leu Asp Ala Ser Gln Glu Lys Val	4798
tca gct ttg acc acc cag ctg aca gca aat gct cac ctt gtg gct gcc Ser Ala Leu Thr Thr Gln Leu Thr Ala Asn Ala His Leu Val Ala Ala	4846
ttt gaa cag agt ctt ggt aac atg aca atc agg ctc cag agt ctg acc Phe Glu Gln Ser Leu Gly Asn Met Thr Ile Arg Leu Gln Ser Leu Thr	4894
atg aca gct gag cag aag gat tca gaa ctg aat gag tta aga aaa acc Met Thr Ala Glu Gln Lys Asp Ser Glu Leu Asn Glu Leu Arg Lys Thr	4942
att gag ctg cta aag aaa cag aac gca gct gcc cag gct gcc att aat Ile Glu Leu Leu Lys Gln Asn Ala Ala Gln Ala Ile Asn	4990
gga gta att aac aca cct gag ctc aac tgc aaa gga aac ggc act gcc Gly Val Ile Asn Thr Pro Glu Leu Asn Cys Lys Gly Asn Gly Thr Ala	5038
cag tct gca gac ctc cgc atc cgc agg cag cac tcc tca gac agc gtc Gln Ser Ala Asp Leu Arg Ile Arg Arg Gln His Ser Ser Asp Ser Val	5086
tcc agc atc aac agt gcc acc agc cac tcc agt gtg ggc agc aac ata Ser Ser Ile Asn Ser Ala Thr Ser His Ser Val Gly Ser Asn Ile	5134
gag agt gac tca aag aag aag agg aag aac tgg gtc aat gag tta Glu Ser Asp Ser Lys Lys Lys Arg Lys Asn Trp Val Asn Glu Leu	5182
cgc agc tcc ttc aag caa gct ttc ggg aag aag aag tcc cca aaa tct Arg Ser Ser Phe Lys Gln Ala Phe Gly Lys Lys Ser Pro Lys Ser	5230

gcg tcc tct cat tca gat att gag gag atg acg gat tct tct ttg cct	5278
Ala Ser Ser His Ser Asp Ile Glu Glu Met Thr Asp Ser Ser Leu Pro	
tcc tca cca aag tta cca cac aat ggg tcc aca ggt tcc acc cca ctg	5326
Ser Ser Pro Lys Leu Pro His Asn Gly Ser Thr Gly Ser Thr Pro Leu	
ctg agg aat tct cac tcc aac tct cta att tca gaa tgc atg gat agt	5374
Leu Arg Asn Ser His Ser Asn Ser Leu Ile Ser Glu Cys Met Asp Ser	
gaa gct gag acc gtc atg cag ctc cga aat gag tta aga gac aag gag	5422
Glu Ala Glu Thr Val Met Gln Leu Arg Asn Glu Leu Arg Asp Lys Glu	
atg aag ctg aca gat atc cgc tta gaa gct ctc agt tct gcc cac cag	5470
Met Lys Leu Thr Asp Ile Arg Leu Glu Ala Leu Ser Ser Ala His Gln	
ctg gac cag ctc cgg gag gcc atg aac agg atg cag agt gaa ata gag	5518
Leu Asp Gln Leu Arg Glu Ala Met Asn Arg Met Gln Ser Glu Ile Glu	
aag ctg aaa gct gag aat gat cgg ctg aag tca gag tct caa ggc agt	5566
Lys Leu Lys Ala Glu Asn Asp Arg Leu Lys Ser Glu Ser Gln Gly Ser	
ggc tgc agc cgg gct cct tcc caa gtg tcc atc tct gcc tcc cgg agg	5614
Gly Cys Ser Arg Ala Pro Ser Gln Val Ser Ile Ser Ala Ser Pro Arg	
cag tcc atg ggc ctc tcc cag cac agc ttg aac ctc act gag tca acc	5662
Gln Ser Met Gly Leu Ser Gln His Ser Leu Asn Leu Thr Glu Ser Thr	
agc ctg gac atg ttg ctg gat gac act ggt gaa tgc tcg gct cgg aag	5710
Ser Leu Asp Met Leu Leu Asp Asp Thr Gly Glu Cys Ser Ala Arg Lys	
gaa gga ggc agg cat gtt aag ata gtt gtc agc ttt cag gag gaa atg	5758
Glu Gly Gly Arg His Val Lys Ile Val Val Ser Phe Gln Glu Glu Met	
aag tgg aag gag gat tcc aga cca cat ctc ttt ctt att ggc tgc att	5806
Lys Trp Lys Glu Asp Ser Arg Pro His Leu Phe Leu Ile Gly Cys Ile	
gga gtt agt ggc aag acg aag tgg gat gtg ctc gat ggg gtg gtt aga	5854
Gly Val Ser Gly Lys Thr Lys Trp Asp Val Leu Asp Gly Val Val Arg	
cggtc ttc aaa gaa tac atc att cat gtc gac cca gtg agt cag cta	5902
Arg Leu Phe Lys Glu Tyr Ile Ile His Val Asp Pro Val Ser Gln Leu	
ggg ctg aat tca gac agc gtt ctt ggc tac agc att gga gaa atc aag	5950
Gly Leu Asn Ser Asp Ser Val Leu Gly Tyr Ser Ile Gly Glu Ile Lys	

cgc agc aac act tcc gaa aca ccg gag ctg ctt cct tgt ggc tat ctg Arg Ser Asn Thr Ser Glu Thr Pro Glu Leu Leu Pro Cys Gly Tyr Leu	5998
gtt gga gag aac acg acc atc tca gtg act gtg aaa ggg ctc gca gaa Val Gly Glu Asn Thr Thr Ile Ser Val Thr Val Lys Gly Leu Ala Glu	6046
aac agc ctg gac tca ctg gtg ttt gag tcc ttg att ccc aag ccc atc Asn Ser Leu Asp Ser Leu Val Phe Glu Ser Leu Ile Pro Lys Pro Ile	6094
ctg cag cgc tac gtc tcc ctc ctg ata gag cac cgt cgg atc att ctc Leu Gln Arg Tyr Val Ser Leu Leu Ile Glu His Arg Arg Ile Ile Leu	6142
tct ggc ccc agc ggc act ggg aaa acc tac ctg gcc aac cgg ctg tct Ser Gly Pro Ser Gly Thr Gly Lys Thr Tyr Leu Ala Asn Arg Leu Ser	6190
gag tat ata gtg ctt cga gag gga cgg gag ttg aca gac ggg gtt atc Glu Tyr Ile Val Leu Arg Glu Gly Arg Glu Leu Thr Asp Gly Val Ile	6238
gcc acc ttt aac gtg gac cat aag tcc agc aag gaa ttg cgc cag tac Ala Thr Phe Asn Val Asp His Lys Ser Ser Lys Glu Leu Arg Gln Tyr	6286
ctg tcc aac ctt gct gac cag tgc aac agt gag aac aat gct gtg gac Leu Ser Asn Leu Ala Asp Gln Cys Asn Ser Glu Asn Asn Ala Val Asp	6334
atg ccc ctc gtc atc atc ctg gac aac cta cac cac gtg agc tct ctg Met Pro Leu Val Ile Ile Leu Asp Asn Leu His His Val Ser Ser Leu	6382
ggc gag atc ttc aat ggg ctg ctc aac tgc aag tac cac aaa tgc cct Gly Glu Ile Phe Asn Gly Leu Leu Asn Cys Lys Tyr His Lys Cys Pro	6430
tac ata att ggc aca atg aac cag gct acc tct tcg act ccc aac ctg Tyr Ile Ile Gly Thr Met Asn Gln Ala Thr Ser Ser Thr Pro Asn Leu	6478
cag ctt cac cat aac ttc aga tgg gtg ctt tgt gcc aac cac acg gag Gln Leu His His Asn Phe Arg Trp Val Leu Cys Ala Asn His Thr Glu	6526
cct gtg aag ggt ttc ctt ggc cga ttc ctg agg agg aag ctc atg gaa Pro Val Lys Gly Phe Leu Gly Arg Phe Leu Arg Arg Lys Leu Met Glu	6574
aca gag atc agt ggg cgg gtg cgc aat atg gag ctg gta aaa atc att Thr Glu Ile Ser Gly Arg Val Arg Asn Met Glu Leu Val Lys Ile Ile	6622

gac tgg att ccc aag gtc tgg cat cac ctc aac cgc ttc ctg gag gct	6670
Asp Trp Ile Pro Lys Val Trp His His Leu Asn Arg Phe Leu Glu Ala	
cac agt tcc tcg gac gtc acc atc ggc ccc cgg ctc ttc ctg tca tgc	6718
His Ser Ser Ser Asp Val Thr Ile Gly Pro Arg Leu Phe Leu Ser Cys	
ccc atc gat gtg gac ggc tcg aga gtg tgg ttc acc gac ttg tgg aac	6766
Pro Ile Asp Val Asp Gly Ser Arg Val Trp Phe Thr Asp Leu Trp Asn	
tat tcc att atc ccc tat ctc ctg gaa gcc gtc aga gaa gga ctc cag	6814
Tyr Ser Ile Ile Pro Tyr Leu Leu Glu Ala Val Arg Glu Gly Leu Gln	
ctc tat gga agg cgc gcc ccc tgg gag gat cct gcc aag tgg gtg atg	6862
Leu Tyr Gly Arg Ala Pro Trp Glu Asp Pro Ala Lys Trp Val Met	
gac aca tat cca tgg gca gcc agc cca caa cag cac gag tgg cct ccc	6910
Asp Thr Tyr Pro Trp Ala Ala Ser Pro Gln Gln His Glu Trp Pro Pro	
ctg ctg cag tta cgg cct gag gat gtc ggc ttc gac ggc tac tcc atg	6958
Leu Leu Gln Leu Arg Pro Glu Asp Val Gly Phe Asp Gly Tyr Ser Met	
cct cgg gag gga tcg aca agc aag cag atg ccc ccc agt gat gct gaa	7006
Pro Arg Glu Gly Ser Thr Ser Lys Gln Met Pro Pro Ser Asp Ala Glu	
ggg gac ccg ctg atg aac atg ctg atg agg ctg cag gag gca gcc aac	7054
Gly Asp Pro Leu Met Asn Met Leu Met Arg Leu Gln Glu Ala Ala Asn	
tac tcc agc ccc cag agc tat gac agc gac tcc aac agc aac agc cat	7102
Tyr Ser Ser Pro Gln Ser Tyr Asp Ser Asp Ser Asn Ser Asn His	
cac gat gac atc ttg gac tcc tct ttg gag tcc act ctg tgacaggggc	7151
His Asp Asp Ile Leu Asp Ser Ser Leu Glu Ser Thr Leu	
ccggagccca gcgccctcct cttctcctca ccgcattcca cctgcattccc ccacatcacc	7211
ctgaagatga cttcctgagc cagccccag ccacagcctt agagctgcgg gaacaccgag	7271
accccccgtc cttcagcctc gacctgggtg caggcatccc gggccagctg cctgcggacc	7331
gcttccttcc acagcgagaa ctgcactacc ttctgttgta cttaattat tgttttgcct	7391
tgttgctgtg acctccctaa gacactgaag atacttctcg ggaaaggatc atgcgggttg	7451
aaatgaaaag agagacagag agagaaaaaa aaaagagaac ccacatgaag ctctgaaacc	7511

aaacagcatc ctgccatgag cttcccagag acagaagaga ctggagcaaa gtcggaaaca	7571
cagagaagca cggcttcccc tcagcacaga ccctccagac tgggtctcag agccgtgcca	7631
cccacccctcc cacacagccg gccacaggga gaactggtgc taaccagggt gcttgcttg	7691
gtcacgttca acgcactaca gagctacgac acaggggaac ctaggagca aataaaccgt	7751
gctttcatgt tttttaaaaaa aaaaaaaaaaa aa	7783

<210> 27  
 <211> 2380  
 <212> PRT  
 <213> Homo sapiens  
 <400> 27

Asn Gln Pro Glu Arg Leu Asn Ser Gln Val Leu Gln Gly Leu Gln Glu	
1 5 10 15	
Pro Ala Gly Glu Gly Leu Pro Leu Arg Lys Ser Gly Ser Val Glu Asn	
20 25 30	
Gly Phe Asp Thr Gln Ile Tyr Thr Asp Trp Ala Asn His Tyr Leu Ala	
35 40 45	
Lys Ser Gly His Lys Arg Leu Ile Arg Asp Leu Gln Gln Asp Val Thr	
50 55 60	
Asp Gly Val Leu Leu Ala Gln Ile Ile Gln Val Val Ala Asn Glu Lys	
65 70 75 80	
Ile Glu Asp Ile Asn Gly Cys Pro Lys Asn Arg Ser Gln Met Ile Glu	
85 90 95	
Asn Ile Asp Ala Cys Leu Asn Phe Leu Ala Ala Lys Gly Ile Asn Ile	
100 105 110	
Gln Gly Leu Ser Ala Glu Glu Ile Arg Asn Gly Asn Leu Lys Ala Ile	
115 120 125	
Leu Gly Leu Phe Phe Ser Leu Ser Arg Tyr Lys Gln Gln Gln Gln	
130 135 140	
Pro Gln Lys Gln His Leu Ser Ser Pro Leu Pro Pro Ala Val Ser Gln	
145 150 155 160	
Val Ala Gly Ala Pro Ser Gln Cys Gln Ala Gly Thr Pro Gln Gln Gln	
165 170 175	
Val Pro Val Thr Pro Gln Ala Pro Cys Gln Pro His Gln Pro Ala Pro	
180 185 190	
His Gln Gln Ser Lys Ala Gln Ala Glu Met Gln Ser Arg Leu Pro Gly	

195	200	205
Pro Thr Ala Arg Val Ser Ala Ala Gly Ser Glu Ala Lys Thr Arg Gly		
210	215	220
Gly Ser Thr Thr Ala Asn Asn Arg Arg Ser Gln Ser Phe Asn Asn Tyr		
225	230	235
Asp Lys Ser Lys Pro Val Thr Ser Pro Pro Pro Pro Ser Ser His		
245	250	255
Glu Lys Glu Pro Leu Ala Ser Ser Ala Ser Ser His Pro Gly Met Ser		
260	265	270
Asp Asn Ala Pro Ala Ser Leu Glu Ser Gly Ser Ser Ser Thr Pro Thr		
275	280	285
Asn Cys Ser Thr Ser Ser Ala Ile Pro Gln Pro Gly Ala Ala Thr Lys		
290	295	300
Pro Trp Arg Ser Lys Ser Leu Ser Val Lys His Ser Ala Thr Val Ser		
305	310	320
Met Leu Ser Val Lys Pro Pro Gly Pro Glu Ala Pro Arg Pro Thr Pro		
325	330	335
Glu Ala Met Lys Pro Ala Pro Asn Asn Gln Lys Ser Met Leu Glu Lys		
340	345	350
Leu Lys Leu Phe Asn Ser Lys Gly Gly Ser Lys Ala Gly Glu Gly Pro		
355	360	365
Gly Ser Arg Asp Thr Ser Cys Glu Arg Leu Glu Thr Leu Pro Ser Phe		
370	375	380
Glu Glu Ser Glu Glu Leu Glu Ala Ala Ser Arg Met Leu Thr Thr Val		
385	390	400
Gly Pro Ala Ser Ser Ser Pro Lys Ile Ala Leu Lys Gly Ile Ala Gln		
405	410	415
Arg Thr Phe Ser Arg Ala Leu Thr Asn Lys Lys Ser Ser Leu Lys Gly		
420	425	430
Asn Glu Lys Glu Lys Glu Lys Gln Gln Arg Glu Lys Asp Lys Glu Lys		
435	440	445
Ser Lys Asp Leu Ala Lys Arg Ala Ser Val Thr Glu Arg Leu Asp Leu		
450	455	460
Lys Glu Glu Pro Lys Glu Asp Pro Ser Gly Ala Ala Val Pro Glu Met		
465	470	475
Pro Lys Lys Ser Ser Lys Ile Ala Ser Phe Ile Pro Lys Gly Gly Lys		
485	490	495

Leu Asn Ser Ala Lys Lys Glu Pro Met Ala Pro Ser His Ser Gly Ile  
 500 505 510

Pro Lys Pro Gly Met Lys Ser Met Pro Gly Lys Ser Pro Ser Ala Pro  
 515 520 525

Ala Pro Ser Lys Glu Gly Glu Arg Ser Arg Ser Gly Lys Leu Ser Ser  
 530 535 540

Gly Leu Pro Gln Gln Lys Pro Gln Leu Asp Gly Arg His Ser Ser Ser  
 545 550 555 560

Ser Ser Ser Leu Ala Ser Ser Glu Gly Lys Gly Pro Gly Gly Thr Thr  
 565 570 575

Leu Asn His Ser Ile Ser Ser Gln Thr Val Ser Gly Ser Val Gly Thr  
 580 585 590

Thr Gln Thr Thr Gly Ser Asn Thr Val Ser Val Gln Leu Pro Gln Pro  
 595 600 605

Gln Gln Gln Tyr Asn His Pro Asn Thr Ala Thr Val Ala Pro Phe Leu  
 610 615 620

Tyr Arg Ser Gln Thr Asp Thr Glu Gly Asn Val Thr Ala Glu Ser Ser  
 625 630 635 640

Ser Thr Gly Val Ser Val Glu Pro Ser His Phe Thr Lys Thr Gly Gln  
 645 650 655

Pro Ala Leu Glu Glu Leu Thr Gly Glu Asp Pro Glu Ala Arg Arg Leu  
 660 665 670

Arg Thr Val Lys Asn Ile Ala Asp Leu Arg Gln Asn Leu Glu Glu Thr  
 675 680 685

Met Ser Ser Leu Arg Gly Thr Gln Val Thr His Ser Thr Leu Glu Thr  
 690 695 700

Thr Phe Asp Thr Asn Val Thr Thr Glu Met Ser Gly Arg Ser Ile Leu  
 705 710 715 720

Ser Leu Thr Gly Arg Pro Thr Pro Leu Ser Trp Arg Leu Gly Gln Ser  
 725 730 735

Ser Pro Arg Leu Gln Ala Gly Asp Ala Pro Ser Met Gly Asn Gly Tyr  
 740 745 750

Pro Pro Arg Ala Asn Ala Ser Arg Phe Ile Asn Thr Glu Ser Gly Arg  
 755 760 765

Tyr Val Tyr Ser Ala Pro Leu Arg Arg Gln Leu Ala Ser Arg Gly Ser  
 770 775 780

Ser Val Cys His Val Asp Val Ser Asp Lys Ala Gly Asp Glu Met Asp  
 785 790 795 800

Leu Glu Gly Ile Ser Met Asp Ala Pro Gly Tyr Met Ser Asp Gly Asp  
                   805                      810                      815  
 Val Leu Ser Lys Asn Ile Arg Thr Asp Asp Ile Thr Ser Gly Tyr Met  
                   820                      825                      830  
 Thr Asp Gly Gly Leu Gly Leu Tyr Thr Arg Arg Leu Asn Arg Leu Pro  
                   835                      840                      845  
 Asp Gly Met Ala Val Val Arg Glu Thr Leu Gln Arg Asn Thr Ser Leu  
                   850                      855                      860  
 Gly Leu Gly Asp Ala Asp Ser Trp Asp Asp Ser Ser Ser Val Ser Ser  
                   865                      870                      875                      880  
 Gly Ile Ser Asp Thr Ile Asp Asn Leu Ser Thr Asp Asp Ile Asn Thr  
                   885                      890                      895  
 Ser Ser Ser Ile Ser Ser Tyr Ala Asn Thr Pro Ala Ser Ser Arg Lys  
                   900                      905                      910  
 Asn Leu Asp Val Gln Thr Asp Ala Glu Lys His Ser Gln Val Glu Arg  
                   915                      920                      925  
 Asn Ser Leu Trp Ser Gly Asp Asp Val Lys Ser Asp Gly Gly Ser  
                   930                      935                      940  
 Asp Ser Gly Ile Lys Met Glu Pro Gly Ser Lys Trp Arg Arg Asn Pro  
                   945                      950                      955                      960  
 Ser Asp Val Ser Asp Glu Ser Asp Lys Ser Thr Ser Gly Lys Lys Asn  
                   965                      970                      975  
 Pro Val Ile Ser Gln Thr Gly Ser Trp Arg Arg Gly Met Thr Ala Gln  
                   980                      985                      990  
 Val Gly Ile Thr Met Pro Arg Thr Lys Ala Ser Ala Pro Ala Gly Ala  
                   995                      1000                      1005  
 Leu Lys Thr Pro Gly Thr Gly Lys Thr Asp Asp Ala Lys Val Ser Glu  
                   1010                      1015                      1020  
 Lys Gly Arg Leu Ser Pro Lys Ala Ser Gln Val Lys Arg Ser Pro Ser  
                   1025                      1030                      1035                      1040  
 Asp Ala Gly Arg Ser Ser Gly Asp Glu Ser Lys Lys Pro Leu Pro Ser  
                   1045                      1050                      1055  
 Ser Ser Arg Thr Pro Thr Ala Asn Ala Asn Ser Phe Gly Phe Lys Lys  
                   1060                      1065                      1070  
 Gln Ser Gly Ser Ala Thr Gly Leu Ala Met Ile Thr Ala Ser Gly Val  
                   1075                      1080                      1085  
 Thr Val Thr Ser Arg Ser Ala Thr Leu Gly Lys Ile Pro Lys Ser Ser

1090	1095	1100
Ala Leu Val Ser Arg Ser Ala Gly Arg Lys Ser Ser Met Asp Gly Ala		
1105	1110	1115
Gln Asn Gln Asp Asp Gly Tyr Leu Ala Leu Ser Ser Arg Thr Asn Leu		
1125	1130	1135
Gln Tyr Arg Ser Leu Pro Arg Pro Ser Lys Ser Asn Ser Arg Asn Gly		
1140	1145	1150
Ala Gly Asn Arg Ser Ser Thr Ser Ser Ile Asp Ser Asn Ile Ser Ser		
1155	1160	1165
Lys Ser Ala Gly Leu Pro Val Pro Lys Leu Arg Glu Pro Ser Lys Thr		
1170	1175	1180
Ala Leu Gly Ser Ser Leu Pro Gly Leu Val Asn Gln Thr Asp Lys Glu		
1185	1190	1195
Lys Gly Ile Ser Ser Asp Asn Glu Ser Val Ala Ser Cys Asn Ser Val		
1205	1210	1215
Lys Val Asn Pro Ala Ala Gln Pro Val Ser Ser Pro Ala Gln Thr Ser		
1220	1225	1230
Leu Gln Pro Gly Ala Lys Tyr Pro Asp Val Ala Ser Pro Thr Leu Arg		
1235	1240	1245
Arg Leu Phe Gly Gly Lys Pro Thr Lys Gln Val Pro Ile Ala Thr Ala		
1250	1255	1260
Glu Asn Met Lys Asn Ser Val Val Ile Ser Asn Pro His Ala Thr Met		
1265	1270	1275
1280		
Thr Gln Gln Gly Asn Leu Asp Ser Pro Ser Gly Ser Gly Val Leu Ser		
1285	1290	1295
Ser Gly Ser Ser Ser Pro Leu Tyr Ser Lys Asn Val Asp Leu Asn Gln		
1300	1305	1310
Ser Pro Leu Ala Ser Ser Pro Ser Ser Ala His Ser Ala Pro Ser Asn		
1315	1320	1325
Ser Leu Thr Trp Gly Thr Asn Ala Ser Ser Ser Ala Val Ser Lys		
1330	1335	1340
Asp Gly Leu Gly Phe Gln Ser Val Ser Ser Leu His Thr Ser Cys Glu		
1345	1350	1355
1360		
Ser Ile Asp Ile Ser Leu Ser Ser Gly Gly Val Pro Ser His Asn Ser		
1365	1370	1375
Ser Thr Gly Leu Ile Ala Ser Ser Lys Asp Asp Ser Leu Thr Pro Phe		
1380	1385	1390
Val Arg Thr Asn Ser Val Lys Thr Thr Leu Ser Glu Ser Pro Leu Ser		

1395	1400	1405
Ser Pro Ala Ala Ser Pro Lys Phe Cys Arg Ser Thr Leu Pro Arg Lys		
1410	1415	1420
Gln Asp Ser Asp Pro His Leu Asp Arg Asn Thr Leu Pro Lys Lys Gly		
1425	1430	1435
1440		
Leu Arg Tyr Thr Pro Thr Ser Gln Leu Arg Thr Gln Glu Asp Ala Lys		
1445	1450	1455
Glu Trp Leu Arg Ser His Ser Ala Gly Gly Leu Gln Asp Thr Ala Ala		
1460	1465	1470
Asn Ser Pro Phe Ser Ser Gly Ser Ser Val Thr Ser Pro Ser Gly Thr		
1475	1480	1485
Arg Phe Asn Phe Ser Gln Leu Ala Ser Pro Thr Thr Val Thr Gln Met		
1490	1495	1500
Ser Leu Ser Asn Pro Thr Met Leu Arg Thr His Ser Leu Ser Asn Ala		
1505	1510	1515
1520		
Asp Gly Gln Tyr Asp Pro Tyr Thr Asp Ser Arg Phe Arg Asn Ser Ser		
1525	1530	1535
Met Ser Leu Asp Glu Lys Ser Arg Thr Met Ser Arg Ser Gly Ser Phe		
1540	1545	1550
Arg Asp Gly Phe Glu Glu Val His Gly Ser Ser Leu Ser Leu Val Ser		
1555	1560	1565
Ser Thr Ser Ser Val Tyr Ser Thr Pro Glu Glu Lys Cys Gln Ser Glu		
1570	1575	1580
Ile Arg Lys Leu Arg Arg Glu Leu Asp Ala Ser Gln Glu Lys Val Ser		
1585	1590	1595
1600		
Ala Leu Thr Thr Gln Leu Thr Ala Asn Ala His Leu Val Ala Ala Phe		
1605	1610	1615
Glu Gln Ser Leu Gly Asn Met Thr Ile Arg Leu Gln Ser Leu Thr Met		
1620	1625	1630
Thr Ala Glu Gln Lys Asp Ser Glu Leu Asn Glu Leu Arg Lys Thr Ile		
1635	1640	1645
Glu Leu Leu Lys Lys Gln Asn Ala Ala Ala Gln Ala Ala Ile Asn Gly		
1650	1655	1660
Val Ile Asn Thr Pro Glu Leu Asn Cys Lys Gly Asn Gly Thr Ala Gln		
1665	1670	1675
1680		
Ser Ala Asp Leu Arg Ile Arg Arg Gln His Ser Ser Asp Ser Val Ser		
1685	1690	1695

COPPER  
1985-1995

Ser Ile Asn Ser Ala Thr Ser His Ser Ser Val Gly Ser Asn Ile Glu  
 1700 1705 1710

Ser Asp Ser Lys Lys Lys Arg Lys Asn Trp Val Asn Glu Leu Arg  
 1715 1720 1725

Ser Ser Phe Lys Gln Ala Phe Gly Lys Lys Ser Pro Lys Ser Ala  
 1730 1735 1740

Ser Ser His Ser Asp Ile Glu Glu Met Thr Asp Ser Ser Leu Pro Ser  
 1745 1750 1755 1760

Ser Pro Lys Leu Pro His Asn Gly Ser Thr Gly Ser Thr Pro Leu Leu  
 1765 1770 1775

Arg Asn Ser His Ser Asn Ser Leu Ile Ser Glu Cys Met Asp Ser Glu  
 1780 1785 1790

Ala Glu Thr Val Met Gln Leu Arg Asn Glu Leu Arg Asp Lys Glu Met  
 1795 1800 1805

Lys Leu Thr Asp Ile Arg Leu Glu Ala Leu Ser Ser Ala His Gln Leu  
 1810 1815 1820

Asp Gln Leu Arg Glu Ala Met Asn Arg Met Gln Ser Glu Ile Glu Lys  
 1825 1830 1835 1840

Leu Lys Ala Glu Asn Asp Arg Leu Lys Ser Glu Ser Gln Gly Ser Gly  
 1845 1850 1855

Cys Ser Arg Ala Pro Ser Gln Val Ser Ile Ser Ala Ser Pro Arg Gln  
 1860 1865 1870

Ser Met Gly Leu Ser Gln His Ser Leu Asn Leu Thr Glu Ser Thr Ser  
 1875 1880 1885

Leu Asp Met Leu Leu Asp Asp Thr Gly Glu Cys Ser Ala Arg Lys Glu  
 1890 1895 1900

Gly Gly Arg His Val Lys Ile Val Val Ser Phe Gln Glu Glu Met Lys  
 1905 1910 1915 1920

Trp Lys Glu Asp Ser Arg Pro His Leu Phe Leu Ile Gly Cys Ile Gly  
 1925 1930 1935

Val Ser Gly Lys Thr Lys Trp Asp Val Leu Asp Gly Val Val Arg Arg  
 1940 1945 1950

Leu Phe Lys Glu Tyr Ile Ile His Val Asp Pro Val Ser Gln Leu Gly  
 1955 1960 1965

Leu Asn Ser Asp Ser Val Leu Gly Tyr Ser Ile Gly Glu Ile Lys Arg  
 1970 1975 1980

Ser Asn Thr Ser Glu Thr Pro Glu Leu Leu Pro Cys Gly Tyr Leu Val  
 1985 1990 1995 2000

Gly Glu Asn Thr Thr Ile Ser Val Thr Val Lys Gly Leu Ala Glu Asn  
2005 2010 2015

Ser Leu Asp Ser Leu Val Phe Glu Ser Leu Ile Pro Lys Pro Ile Leu  
2020 2025 2030

Gln Arg Tyr Val Ser Leu Leu Ile Glu His Arg Arg Ile Ile Leu Ser  
2035 2040 2045

Gly Pro Ser Gly Thr Gly Lys Thr Tyr Leu Ala Asn Arg Leu Ser Glu  
2050 2055 2060

Tyr Ile Val Leu Arg Glu Gly Arg Glu Leu Thr Asp Gly Val Ile Ala  
2065 2070 2075 2080

Thr Phe Asn Val Asp His Lys Ser Ser Lys Glu Leu Arg Gln Tyr Leu  
2085 2090 2095

Ser Asn Leu Ala Asp Gln Cys Asn Ser Glu Asn Asn Ala Val Asp Met  
2100 2105 2110

Pro Leu Val Ile Ile Asp Asn Leu His His Val Ser Ser Leu Gly  
2115 2120 2125

Glu Ile Phe Asn Gly Leu Leu Asn Cys Lys Tyr His Lys Cys Pro Tyr  
2130 2135 2140

Ile Ile Gly Thr Met Asn Gln Ala Thr Ser Ser Thr Pro Asn Leu Gln  
2145 2150 2155 2160

Leu His His Asn Phe Arg Trp Val Leu Cys Ala Asn His Thr Glu Pro  
2165 2170 2175

Val Lys Gly Phe Leu Gly Arg Phe Leu Arg Arg Lys Leu Met Glu Thr  
2180 2185 2190

Glu Ile Ser Gly Arg Val Arg Asn Met Glu Leu Val Lys Ile Ile Asp  
2195 2200 2205

Trp Ile Pro Lys Val Trp His His Leu Asn Arg Phe Leu Glu Ala His  
2210 2215 2220

Ser Ser Ser Asp Val Thr Ile Gly Pro Arg Leu Phe Leu Ser Cys Pro  
2225 2230 2235 2240

Ile Asp Val Asp Gly Ser Arg Val Trp Phe Thr Asp Leu Trp Asn Tyr  
2245 2250 2255

Ser Ile Ile Pro Tyr Leu Leu Glu Ala Val Arg Glu Gly Leu Gln Leu  
2260 2265 2270

Tyr Gly Arg Arg Ala Pro Trp Glu Asp Pro Ala Lys Trp Val Met Asp  
2275 2280 2285

Thr Tyr Pro Trp Ala Ala Ser Pro Gln Gln His Glu Trp Pro Pro Leu

2290	2295	2300
Leu Gln Leu Arg Pro Glu Asp Val Gly Phe Asp Gly Tyr Ser Met Pro		
2305	2310	2315
Arg Glu Gly Ser Thr Ser Lys Gln Met Pro Pro Ser Asp Ala Glu Gly		
2325	2330	2335
Asp Pro Leu Met Asn Met Leu Met Arg Leu Gln Glu Ala Ala Asn Tyr		
2340	2345	2350
Ser Ser Pro Gln Ser Tyr Asp Ser Asp Ser Asn Ser Asn Ser His His		
2355	2360	2365
Asp Asp Ile Leu Asp Ser Ser Leu Glu Ser Thr Leu		
2370	2375	2380

<210> 28  
<211> 96  
<212> DNA  
<213> Homo sapiens

<220>  
<221> CDS  
<222> (1)..(93)

<400> 28

agc agg gag agg gga ggg agt gtg ccg tct ctt ctg caa ggg cag tgc	48
Ser Arg Glu Arg Gly Gly Ser Val Pro Ser Leu Leu Gln Gly Gln Cys	
ccc agc ctc agc cac act tct gat ctg cag tcc aac aga cct ttc tag	96
Pro Ser Leu Ser His Thr Ser Asp Leu Gln Ser Asn Arg Pro Phe *	

<210> 29  
<211> 75  
<212> DNA  
<213> Homo sapiens

<220>  
<221> CDS  
<222> (1)..(72)

<400> 29

cat gcc aaa gag aac ctg ggg gtg cca ggg ggt cct cag agc tca cac	48
His Ala Lys Glu Asn Leu Gly Val Pro Gly Gly Pro Gln Ser Ser His	

tgc act tgt ggc acc cac agc gag tag	75
Cys Thr Cys Gly Thr His Ser Glu *	

<210> 30  
 <211> 489  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> CDS  
 <222> (1)..(489)

<400> 30

cca tcc gtg agc cga gga aac tgt aca cag atc tac aca gac tgg gcc	48
Pro Ser Val Ser Arg Gly Asn Cys Thr Gln Ile Tyr Thr Asp Trp Ala	
aat cat tac cta gcc aaa tcc ggc cac aag cgt ctc atc aag gat ctc	96
Asn His Tyr Leu Ala Lys Ser Gly His Lys Arg Leu Ile Lys Asp Leu	
cag caa gat gtg aca gat ggc gtc ctc ctg gcc cag att atc cag gtt	144
Gln Gln Asp Val Thr Asp Gly Val Leu Ala Gln Ile Ile Gln Val	
gtg gca aat gaa aag att gaa gac atc aat ggc tgt ccg aag aac aga	192
Val Ala Asn Glu Lys Ile Glu Asp Ile Asn Gly Cys Pro Lys Asn Arg	
tcc caa atg att gaa aac ata gat gcc tgc ttg aat ttc ctg gca gct	240
Ser Gln Met Ile Glu Asn Ile Asp Ala Cys Leu Asn Phe Leu Ala Ala	
aag gga ata aac atc cag ggg ctg tct gca gaa gag atc agg aat gga	288
Lys Gly Ile Asn Ile Gln Gly Leu Ser Ala Glu Ile Arg Asn Gly	
aac ctc aag gcc att cta ggc ctc ttc agc ctc tcc cga tac aag	336
Asn Leu Lys Ala Ile Leu Gly Leu Phe Phe Ser Leu Ser Arg Tyr Lys	
cag cag cag cag ccc cag aag cag cac ctc tcc tca cct ctg ccg	384
Gln Gln Gln Gln Pro Gln Lys Gln His Leu Ser Ser Pro Leu Pro	
ccc gcc gta tcc cag gtg gcc ggg gcc ccc tcc cag tgc cag gct ggc	432
Pro Ala Val Ser Gln Val Ala Gly Ala Pro Ser Gln Cys Gln Ala Gly	
acc cct cag cag cag gtg cca gtc act ccccaa gcc ccg tgc cag cct	480
Thr Pro Gln Gln Val Pro Val Thr Pro Gln Ala Pro Cys Gln Pro	
cac cag cca	489
His Gln Pro	

<210> 31  
<211> 31  
<212> PRT  
<213> Homo sapiens

<400> 31

Ser Arg Glu Arg Gly Gly Ser Val Pro Ser Leu Leu Gln Gly Gln Cys  
1 5 10 15

Pro Ser Leu Ser His Thr Ser Asp Leu Gln Ser Asn Arg Pro Phe  
20 25 30

<210> 32  
<211> 24  
<212> PRT  
<213> Homo sapiens

<400> 32

His Ala Lys Glu Asn Leu Gly Val Pro Gly Gly Pro Gln Ser Ser His  
5 10 15

Cys Thr Cys Gly Thr His Ser Glu  
20

<210> 33  
<211> 163  
<212> PRT  
<213> Homo sapiens

<400> 33

Pro Ser Val Ser Arg Gly Asn Cys Thr Gln Ile Tyr Thr Asp Trp Ala  
5 10 15

Asn His Tyr Leu Ala Lys Ser Gly His Lys Arg Leu Ile Lys Asp Leu  
20 25 30

Gln Gln Asp Val Thr Asp Gly Val Leu Leu Ala Gln Ile Ile Gln Val  
35 40 45

Val Ala Asn Glu Lys Ile Glu Asp Ile Asn Gly Cys Pro Lys Asn Arg  
50 55 60

Ser Gln Met Ile Glu Asn Ile Asp Ala Cys Leu Asn Phe Leu Ala Ala

65	70	75	80
Lys Gly Ile Asn Ile Gln Gly Leu Ser Ala Glu Glu Ile Arg Asn Gly			
85	90	95	
Asn Leu Lys Ala Ile Leu Gly Leu Phe Phe Ser Leu Ser Arg Tyr Lys			
100	105	110	
Gln Gln Gln Gln Pro Gln Lys Gln His Leu Ser Ser Pro Leu Pro			
115	120	125	
Pro Ala Val Ser Gln Val Ala Gly Ala Pro Ser Gln Cys Gln Ala Gly			
130	135	140	
Thr Pro Gln Gln Gln Val Pro Val Thr Pro Gln Ala Pro Cys Gln Pro			
145	150	155	160
His Gln Pro			

<210> 34  
<211> 479  
<212> DNA  
<213> mouse

<220>  
<221> CDS  
<222> (3) ... (476)

<400> 34

at gag aag agc cga aca atg agt cgg tca ggc tcc ttc cgg gat ggg 47  
Glu Lys Ser Arg Thr Met Ser Arg Ser Gly Ser Phe Arg Asp Gly

ttt gag gaa gtt cat gga tcc tcc ctg tcc ttg gtt tcc agc aca tcc 95  
Phe Glu Glu Val His Gly Ser Ser Leu Ser Leu Val Ser Ser Thr Ser

tcc atc tac tcc acg cca gaa gaa aaa tgc cag tca gag att cga aag 143  
Ser Ile Tyr Ser Thr Pro Glu Glu Lys Cys Gln Ser Glu Ile Arg Lys

ctg agg cga gaa ctg gat gcc tcc cag gaa aag gtg tct gcg ctg act 191  
Leu Arg Arg Glu Leu Asp Ala Ser Gln Glu Lys Val Ser Ala Leu Thr

acc cag ctg act gca aat gct cac ctt gtg gca gcc ttc gag cag agt 239  
Thr Gln Leu Thr Ala Asn Ala His Leu Val Ala Ala Phe Glu Gln Ser

ctg gga aac atg acc atc agg cta cag agt tta act atg acc gct gag 287  
Leu Gly Asn Met Thr Ile Arg Leu Gln Ser Leu Thr Met Thr Ala Glu

cag aag gat tca gaa ctg aac gag tta aga aaa acc atc gag ctg ctg 335

Gln Lys Asp Ser Glu Leu Asn Glu Leu Arg Lys Thr Ile Glu Leu Leu	
aag aaa cag aat gca gct gcc cag gct gcc att aat gga gtg att aac	383
Lys Lys Gln Asn Ala Ala Gln Ala Ala Ile Asn Gly Val Ile Asn	
acg cca gag ctc aac tgc aaa gga aat ggc agt gcc agg cta cag acc	431
Thr Pro Glu Leu Asn Cys Lys Gly Asn Gly Ser Ala Arg Leu Gln Thr	
tac gca tcc gca gca aca ctc ctc cga cag tgt ctc cag tat caa tag	479
Tyr Ala Ser Ala Ala Thr Leu Leu Arg Gln Cys Leu Gln Tyr Gln *	
<210> 35	
<211> 22	
<212> DNA	
<213> mouse	
<220>	
<221> CDS	
<222> (1)...(21)	
<400> 35	
cgc cac cag cca ctc aag tgt g	22
Arg His Gln Pro Leu Lys Cys	
<210> 36	
<211> 158	
<212> PRT	
<213> mouse	
<400> 36	
Glu Lys Ser Arg Thr Met Ser Arg Ser Gly Ser Phe Arg Asp Gly Phe	
1 5 10 15	
Glu Glu Val His Gly Ser Ser Leu Ser Leu Val Ser Ser Thr Ser Ser	
20 25 30	
Ile Tyr Ser Thr Pro Glu Glu Lys Cys Gln Ser Glu Ile Arg Lys Leu	
35 40 45	
Arg Arg Glu Leu Asp Ala Ser Gln Glu Lys Val Ser Ala Leu Thr Thr	
50 55 60	
Gln Leu Thr Ala Asn Ala His Leu Val Ala Ala Phe Glu Gln Ser Leu	
65 70 75 80	
Gly Asn Met Thr Ile Arg Leu Gln Ser Leu Thr Met Thr Ala Glu Gln	
85 90 95	

Lys Asp Ser Glu Leu Asn Glu Leu Arg Lys Thr Ile Glu Leu Leu Lys  
100 105 110

Lys Gln Asn Ala Ala Ala Gln Ala Ala Ile Asn Gly Val Ile Asn Thr  
115 120 125

Pro Glu Leu Asn Cys Lys Gly Asn Gly Ser Ala Arg Leu Gln Thr Tyr  
130 135 140

Ala Ser Ala Ala Thr Leu Leu Arg Gln Cys Leu Gln Tyr Gln  
145 150 155

<210> 37

<211> 7

<212> PRT

<213> mouse

<400> 37

Arg His Gln Pro Leu Lys Cys  
5

<210> 38

<211> 19

<212> PRT

<213> artificial Sequence

<220>

<223> Description of the artificial Sequence: Peptide for production of AK

<400> 38

Glu Lys Gly Glu Asp Pro Glu Thr Arg Arg Met Arg Thr Val Lys Asn  
5 10 15

Ile Ala Asp

<210> 39

<211> 81369

<213> DNA

<214> Homo sapiens

<400> 39

gatcagactt tgaagagtgt ttgtaccatg ctaaagttta cagaatttat tcctgcttt 60  
tgagggtgca ttgcaaatcc aggctagagg gagagatacc agttaggana gtacagcaat 120

actctactgg	gaaatggtga	ggtgttcgt	gaagacaatg	gcaacacaga	tgaagacatg	180
cagatggagg	aaataaaagat	ccagttgagc	ttgttgcca	gttggataga	ggtttaggtt	240
atgcatgatg	gagcaatcta	ggttttgtc	ttgggttaggt	gttccatga	tagtactcag	300
aatgaatcat	atagttgtac	aggttgaatc	ccaccatgt	ttgcacaata	gagtgactgt	360
ctagctgaaa	tccagatgac	actctgtatg	ctaagctatg	cttcatggaa	ctgtataaag	420
gcacttgcta	cataggctag	tggcagatct	ggaagtaacc	tatatggtat	ataggaaatg	480
aggtggctt	tgtataaatac	ctacagataa	atttcatttc	ctgatcctat	tattttgact	540
catgttagcc	caagaagagt	attcagtaact	tcatatccct	gaaggtaaga	cagagtagta	600
ttagattcac	tatttggcaa	ataaaaaggga	tcaagtccct	agatcaagct	gatgaatcaa	660
cacccatag	gatatgtccc	aaccaattat	atggcttccc	ctataaataa	aatctagttc	720
tcttctctgg	agaggaacag	tgaagaatat	cataacctat	gctacaaact	gcttgagtag	780
gagctacttc	tctccaaaggc	tttatatcat	tcattctggc	aggccccct	gtttgttctc	840
accagctcct	gggaaattta	tttctcctct	agtgatataa	aagctctctg	tttgagatga	900
agggctgccc	agtttatcag	atctgtatta	gtctgttctc	aggctgctaa	taaagacata	960
cctgagactg	agtaatttat	gaaggaaaga	ggttaattt	actcacagtt	ccacatggct	1020
ggggaggcct	cacaatcatg	gcaaagact	aataaggagc	aaagtcacat	cttacatggc	1080
tgccagacaag	agagcatgtg	cagggaaact	gctctccata	aaaccatcag	atcttgtgag	1140
acttgttcac	tattacaaga	acaacagaca	ggaaaacccg	ccccctcaat	tcaattacct	1200
gccactggga	ccctcccaca	acacatgggg	attatgagag	ctacaattca	agatgagatt	1260
tgggtgggaa	taccgccaaa	ccatatgaag	ttcttcttt	gttactgggt	accatatcca	1320
ttctgttgag	gttctgagcc	tttccagtt	ctgtaactcc	tctatctcct	gtctgtgcta	1380
agactcagtg	acctctctct	gccttgcttc	tgcttgc	tgacccttcc	tgtgcacat	1440
ctcactctag	tttgcacc	tgaggtgaga	gatggccag	attagcaaca	acaatctgt	1500
gactaaaatc	ctcttttaggg	aggaagcaaa	attcagatgg	atgttactaa	acaagctca	1560
gaaacagaga	ccaggggtgt	ggaagtaagg	tagtagcctg	agagcagctg	gcagtgttt	1620
agacctggag	ggaggttagg	tcatcagcaa	tgaggagact	gcctggaaaa	tcctagaaaa	1680
ttaagacatc	tggtcaggca	aggtcatatc	accagcacac	ttccctttc	aagttgaatc	1740
ccttcctct	gttaagagga	ttcaagtgtc	tttcttgcat	tttgtcttct	cttctatatc	1800

catgcttgca atataaggag acagcagttg gctgttggc ctagaaaata taaatggcca	1860
ttttgaaagc atgccagaca ggatctgcgg caagtttca atgttactgc tgccatctgt	1920
tgttcttcag tgctggatg tgaatctctt ggcaaacatc tctctaattc tgaactatct	1980
ttcacccccca tctagagata ttcaacttact gaagtgcctt tttaaagcaa tgttcctcac	2040
caaggcgatg ttctgaatgt tttaaaatgg aagaatctgg aatgtttta ttataataca	2100
ttttgtatat cccaaagcaa aaatcaattt ctcatggtt aatacttttga taattttgtt	2160
tttaataata ttttcctttt aaatataaga aatattttat tgaattaata cttaatgtt	2220
gctgttcaa gtaagataaa acagaacaga ttactgttt caaccctgtt cacagttgc	2280
tctgttaacta agttgtttag ctttatctaa gctttttat ttacataaa cgttccctt	2340
ttcacttaac cttgaaattta tagtaatttgc ggaacttcta ttccctctgaa agagaaagct	2400
aatgccaaag atatttcaag ggagaaaagaa ggttttaaa aggagagaca attcagctca	2460
gacttaatag ctgtgattgc tatttattaa gcagaacgcc tataactaaa ttctcagata	2520
tccaaaaaac agcctgtaca ttctcaaaag tgaagattac acattttcta agttaaggta	2580
aaagttttgt ctctgttagca tcttactgat ttctatcttc tcattctgcc ttaataatgt	2640
cactaaataa atgtttgatg cactaataca tgaataaaac tattcatggt aatgattttt	2700
tagaaacaca gctaagttt gtaattttgt tttaaaaaaa ttaaaaattt aaatataaaaa	2760
atgttttaa aaggcttgaa ttcttgtaa aatgtacaca tttaagttg taggctgtct	2820
ttaaaaataa tctctccaca cactgttagta tttaaaacat catgatattt ctataaaaca	2880
tcaacaaataa gggcagtgga aaacatggta atcactaaaa atgctcacat gtcataattt	2940
aagacttgat aagtaaacca caataataaa tagaaaagaa atagttgtct aaaaagggat	3000
tctcacctt caaaccttac cataaaaatg gaatataaaa gaaggaagag gaggagaaat	3060
caaattatat cataaaaattt tctggcaaa aatattacag aagaaaataa gaaagattta	3120
tggagttgac taaaacattt ttgaatccta tacataaaaa tatcgtaat taaaaggaaa	3180
aacaaagaaa cagatttggg aaatatttga aactggttt tttagcat ttaaaaatgt	3240
aatacaaataatg gattattttaa actccattgc aaaaatacac aaaggacatt gacaatgtct	3300
gaaaataaaa ttagctaagt aagttataga aaaactcagt ctcacaattt gacaaatgtt	3360
actgaaaact attaatataa ttagtaacta ttacatg tcaaaaatttt tgaattacta	3420
aaggaaacca caatgcctga aagtatccag ggtttttttt ttttttataa atattggcac	3480
tgtcatatgg gtggcaggaa ttgaagtgtat gttgtttctt cagttattaa gttgcattctg	3540

cagtgtttca aatgtccaaa acctgtgagt cagtaattct ctttttgtat atttaccca 3600  
ataacaataat tctaaacata atctcaatat atatgtacaa agttattcac tgcatgtta 3660  
cttacaatag tttagaaaatt gtaaaatgct ttatgcacatc taaaatataa attgttgaat 3720  
ataataatgt ccatatgata taattatatac attattataa ataatgaatt agaaaataat 3780  
ttaagagcat taaaataatt ataaggtaat atgaagtgaa tgaataatgt acagatacta 3840  
taatcagcag agtgttaact aggttaattt ttatgtgtgt atataactact tcctaaaaat 3900  
gacttgacag aaatcatcaa aatgctaattt gtggttactt ctgggtggga atacagatga 3960  
tttactttgt tcctttatg tatttcgtca ctgcccagtc ttccacagtg agcatatatt 4020  
ggtttttaaa tttatataag atggaaaaag ataccaaattt gtcttcaatg aatcctggag 4080  
ttaactttca tgtgtgtcat atgttatatt ctaaacttat cacaaataga agactttaaa 4140  
tcaacttgta cctatttcaa ctatataaca gcattttaa aatgagcatt gaattaaact 4200  
accaaaaacca accatcatga ggattattca agtaatgtgt ttaaacaatggaa 4260  
taaaattact ttatctcattt tgtgatttca gccattttaa aaaaaataga tgtttctact 4320  
ctccttcaga tatcattttaa acataaactt gtgcctgact gcataaatcc cttttaaact 4380  
aatatcactt attacgtttta actaagtcta cctagggtttt cttgtataa agaacaagag 4440  
ctttccattt tttgtttacc tagcccttcc tgatgccacg acagaatagc tgtaaatctt 4500  
cattattttt attctagaga aaataaaacccca acccacatgg tattaatgtt agtctctatg aatatttcat 4560  
ggttgttctc tttactcaaa acccacatgg tattaatgtt agtctctatg aatatttcat 4620  
ggataaaaatc agagcattaa gtgcataacta aaaacaataa gaatggaaag actttaccc 4680  
tatgtttata tgaatttcta ggttatcaag aagtttataag gctataggct ataaagtctt 4740  
aggctatgat atagtaacct aatgttagact tcccttgcata catgaaaata atggtagttaa 4800  
gtacaaacag aagatgagct taaaattatt ctttgagtcc tcttgatggaa tttttcccc 4860  
cacactttcc cccaaattgt tttatgccta tattgttagga gaccatgcaa gagacctaga 4920  
gtctctttt ctttcatcac tttccaaatca acagcaaatc ctatcatttt taccacaaaa 4980  
tatatcttga aactcccttc ttttgatttca cttgtactc cccatcaaaa actgaagagt 5040  
gtcacaatac ttcatcaatgt tccctacttg cactctaccc ttaatataatt tgtagcacta 5100  
aaatgtttt aaaaatataa tctgctttagt tcattttact gctcaataact atctgatttt 5160  
ctattqcact tctaaqatac tctaatttct tagcactctca tataaaaatcc ttaaqqqqct 5220

tccctgctca cctttcaga ctcagaacta tgtatttcct tttgcctgct gtacttgta	5280
cactggattc ttgatttttgc ttacttccag gttttcacac ttatTTTACAC aataaatgtg	5340
aaataccctt tttgacaata tctacaaata tttcttattt gtctttatttgc ctcttcctg	5400
taatgttttag tcttcatttt cctgataatg gctatctaaa gttatctcct caaagaagca	5460
gttatttttacc cacccaaatac ttcttagtcct tctctggagt tttcttcata ctccattccc	5520
ttggtttttg ccacaatttg taataatttg caatttggag tgtagaaatg agggaaataaa	5580
tcacaggtaa tgactatagt ttgtgactat gtaagattgg attcgatttatttattcc	5640
acaaacactg aggcaactgca tttagccaaa tgccaatctt gggcagttag actctgaaag	5700
agaatctgct tcccccacca taaactacaa agtgaacaca ctcagaatgt acataaatta	5760
cagaatgaaa gcacactaga agtaaacaca gatgtggaag aggtaaatgt tcctgaaaaa	5820
tcatggaaag attcataaaag ggaatgacat ttcaactgga ttctaaacca gttattcaag	5880
ctccacaagg ttgcacagta aatgaggcgt ggcaggatga cataccttag aaagtaaaag	5940
gaatctttt taaactgcta taaaaatcat tacatataca tttttaggt cgagagtaag	6000
gtatTTAACA taaaatcatt ttagtatATC agtgtttata tagacttagg ttttctcat	6060
ttaaaacctc tttaatgac ttgtgtttt cttcatggta ataaaacatt ttcccaggaa	6120
gtgctgaata aatcttctt gaaatacgtt ttattgcttt ctatcaatga ccctgaagta	6180
atacagaatt tacacttcag cgggtgcaat gctcaaactt gacaggtaat gcactgtgtt	6240
tgctgatata agaggtatga tttgggctt agtggTTTG tgctcattta gcttcagga	6300
gaaaataatt gacttaacat tttgatacta aaacccaaag cctaaacagttaatttcttgg	6360
atTTAAATT attattgcaa agattattgt gccgaataat atgaaaatat tttatataat	6420
atTTAAAAG tatatcttt tcttggattt atttaaatta ccataaaaat gtgcggaaaaa	6480
gttataactga aatgtgatag gatctttaa aagtggtgcc ttgattttgt taagtgttac	6540
ctagtttcc tctgaaaaca agaaacatac ccagaagttt tcacgaaatg gtctcatgaa	6600
tatctaaggtagt tagccgttag tctcatctga gacaaggaaa gtccttcca ctatgagcct	6660
gtaaaatcac aagcaagcta gtacttcct agatacaatg ggagtagtgg tattggtaa	6720
acacagctgt ttcaaatggg agaaattggc caaaattaat gggttacagg gcatgcaatt	6780
ccggaaatcca tctggcagt caaattgtaa aactccaaaa tgatntctt tgactccatg	6840
tntcacatcc aggacatgct gangcaagag ataggttccc ataatcttg gcagctctgc	6900
ccctgtggct ttgcaggta tatcaccctt cccagctgct ttcacaggtt ggcatttgagt	6960

gtctgtggct ttcccaggaa caaggtgcaa gctgtggtg gatctaccat tctggggttt	7020
ggaggatgat ggccctcttc tcatacgctcc actaggccgt gctccagtag agactctgtg	7080
ggggctctga ccccagattt ccctcctgca ctgccctagc agagattctt catgagggcc	7140
gtgcccctgc agaaaactct ttcttggca tccaggcatt tccatacatc tgaaatctag	7200
gtggaggttc ccaaaccctcg attcttaatt tctgtgcacc tgcaggctct ctaccacgtg	7260
gaagctgcca aggtttgggg ctgcaccct ctgaaaccac aggctgagct ataccttggc	7320
cccttttagc aatggctgga gtgactggga cacagggcac caagtctcta ggctgcacac	7380
agtatggca ccctggggcc agccctcaaa atcattttt cctccttaggc ttctggatca	7440
gtgaagggtg gggctgccat gaagacctat gacatgccct ggagacattt tccccattgt	7500
cttggggatt aacactggct ccttgttaact tatgcagatt tctgcagcca gctgaatttc	7560
tcctcaaaaa atgggttttt ctttctact gcattgtcag gctgcaaatt ttctgaactt	7620
ttatgctgtt tccctttaa aatgcgatgc tctaacaaca cccgtcacct cttgaatgct	7680
ttgctgctta gaaatttctt ctgtcagata ccctaaatca tctctctcaa gttcagagtt	7740
ccacaaatct cttagggcagg ggcaaaatgc caccagtctc tttgctaaaa cataacaaga	7800
gtcgcccttg ctccagttct cagcaagttc ctcatctcca tccgagacaa cctcagcctg	7860
gtccttattt tttatatcac tataaaaatt tttgtcaaag ccattcaaca agtctctact	7920
ccaaactttc ccacattttc ctgtcttctt ctgagccctc caaattgttc cagcctctgc	7980
ctgatacaca gtcrrrrrrracttccaca tttttggata tctttcagc aatgccccgc	8040
tctactggta ccaacttact ttgttagtcc gttttcacac tggtgataaa gacataccca	8100
agactggaaa gaaaaaaaaagg tttaatttggaa cttacagttc cacatggcta gggaggctc	8160
acaatcatgg caggaggcaa aaggcatttc ttacatgatg gcagcaagag aaaatgagga	8220
agatgcaaacc gcagaaatcc ctgataaaac catggacac tgtaagactt attcactacc	8280
actaggacag tatgggtgat accacccca tgattcaaattt gatctccaac caggtgcctc	8340
ccacaacaca tggaaattt gggaaatcaa ttcaagatga gatttggta gggacacaga	8400
gccaactat atcacatgga tttcttatac ttttgcattt aataacacaa aaaaaaaaaat	8460
acatcattaa aagtttagaa gtgagaaggt gttttatgg aaatcaaaaa taatatcacc	8520
ttagtgaaca gtattctt gattgttagtt gaatttagaga gcagaataca tctagaagat	8580
tcagtagtaa gcatgtttct tcgattaatg gaaaatttga atagcctagc tgattgagat	8640

tgaggttact attaaatgcc tgaagtataa gagttgggtt tttatgtaaa caaaatatct	8700
gaaaaattttacatg tacatgtgta agtaggactg ttgagccccca gtaacatgaa atatcaaaga	8760
gcatgactcg aatacctgcc atatgaagtg ctattacatc aaaaaagagg cgtgtgctga	8820
aaaattacct acaaatggca ttttcctcaa atcaatttta aatcttcaga atttcatttt	8880
aataattgtt tagttaatat ttcagaatcc ctcatcataa aaagcaggca aaaggcaaaa	8940
gtccttgaat gtataacaca tttgtttca aacaagcctg cctctaactg tgaatccagg	9000
agtgaatcca gaactacaaa ttaactaaga ttggcccccattt cgagttactg aacgttaaaa	9060
atctaaaaac taaaaggcat gcctcaacaa ttattttctt cttggaatca ttaattaacc	9120
tatgtgtatc caaacaataa tcttccagca gtttcgctag ctacatTTT aattacttaa	9180
tatcatgtaa aatttggttt attattgttc agttctgaat tttgacatat gcatcaagcc	9240
atgcaactgc taccacagtc ttccctgatca ctgatctgtt ctaaatctct atagcatttt	9300
tcctttctt aaatgttgca taaataaaac cataccttat gtggcctttt gaatctggca	9360
tctttaactt aatgcgccttg aaattaatct atgtcatttc atgtatcaat ggctcaatct	9420
ttttaattgt taagaaaaaa tgtatgctgg gataaaatatc tttctaaatg agttttgtt	9480
cacaatgctg agtgtttgtt taggatagag tccttagaaat ggtatcacta ggtcaaacat	9540
tcaaataatt ttaaaatatt tgatacatat tgccaaataa tctcaaattt tttaccaata	9600
tacatTTTatc acagtatggg ataaatgtgt ctttcttata ccaactgaca acattaatga	9660
taatacataa aatattctt gctaatttga tgggacagaa atgttatatc cttattagca	9720
ttttattttt gtgggtgaat gactgtactg tacagccaga gatatttggt tcaaaatcca	9780
tcttcattat ttactgtatg taaaaattta ggtgagctat ttaatctctt gatgccttag	9840
tctcctaattc tataaagtgg ggataattgt accaatcata ttaggttcct gtgagaatta	9900
actgaattac tatagaaaat gcttagaaatg gtatctagtc accaggaagg actctctctg	9960
tattactgt ttattatcta acacgtttaa ttattaatga agctcagttt cgttatatgc	10020
ttgggatatt tgaaactttt cttagtgaat tttccaataa aattatttgt ctatTTTct	10080
atggacaagt tggtattatt cttagtggt tgTTTcaggt tcagtttagta agaattttaa	10140
ggattttcta tcacattta gcaaactttt tctgcatttt atctttttc tttcagataa	10200
tgtttgcaaa atgtaaaaaa aacaaaagggt ttcttcatca agttggtatc tttatcttt	10260
ttattgcttt gtgatttcaa aattcttgc tcgagaacca aaatataatat ttgatgaaat	10320
agttctcttc ttttactcat tctgaagtca ttggaattga atttggcata tgatataaaat	10380

cctaatttta tattttatga tattcaaaat ttctaacaaa tatttactta ataatcta	10440
ccaggtttct attgtttctt ctgtttcctt tataatgctt tttctgaagt tattttctt	10500
agacttaaat attagtataa tattatcata gagaaaaaaa tatctgttag ctatgaataa	10560
aaggctttca tcttattgtt gcattaatat atttaaatgt agagagcata cagattagca	10620
aagaaaaagt ataattgcct tttttatag ttgacatgaa catgtataaa gaaaaaccaa	10680
aaaaatcaat aaaacaacta gaacttatta gtgaatttag caagatcata gcatacaaag	10740
ccaagattca aaattccatt ttatattatct actaacaaaa aatatttcaa atttgaaaat	10800
ttaaatatgc catttacaat aacatcaaaa tattgaacaa taaagtatTT aggaatttat	10860
aaaatgaaat ctcctataacc aggaattaca gaccattgct gaaataaatg aaagaagacc	10920
aatatatgtg aagagatact catttgcggta ttgagagaca atattgttaa agtacgtat	10980
tttcccaaatt taatcaatag attcaatata atggtaaca gaacaccaga agatgttctg	11040
togaagctga caagctatTT ctataattca aatggaaatg caaaaggcag tcactgcca	11100
caccagcatg gactgtctgg gttccagtag gttacttcac tactgcctct tctgtcagcc	11160
acatcacgac agctgcccg aagccagaga aactcctcac acctggccca ctgctgcagc	11220
taccagcatc caggcaagcc accatcagcc cactggtaac tgccaacaga ggtaccactg	11280
tacactaccc tggggAACAA agataggcat gtagtcagcc cacctctgcc accactaggg	11340
cctgaagcct ggcccacctg acactgcagt cctcagcaca gttcatcac agttctgtt	11400
aataaccaca ccctaaccta ccaaggaaat cacaaatgtc actgacactg tttgtagcca	11460
aagaaatcat agagagacta cattactgca cacaccata atcaaagcca cagtacccta	11520
tccagacaac atcacaggtt tatctaaagg aaaaaatttt cccatatgaa agcgaattca	11580
aatataggaa gaagcgactg ttacaacaga tatgcagata aagttcaac aatatcctac	11640
attcaaccag aagaaagaat ctcagaaggta aagacaggt cttctgaaat aatcttagtca	11700
gacaaaatta aaagagaata atcaaattctt tcctgacatt tggataaca ttAAAGTGC	11760
caaataatacg aattatacgat acccctgaga gtgaaaagac aaagaaaaga tttagaaaacc	11820
cacttaatta aataatataat gaaaacttcc taagtctgc aagagttttt gatatttggg	11880
atgcaggagg ctcaatggtc cccaggccga taaaacgcaa aaaggtctta tacacagcac	11940
attacaatca gactgtttaa agtcaaagat aaggaataaa ttctaaaaac agcaagagaa	12000
agtgtatgtt aacctatgaa gtaaacctta tcagactgac agcaaatttc tggcagaaac	12060

tttacaggcc agaaagaata ggacaatata ttcaaagtgc ttaaagaaaa aaaaaactat	12120
cagcctaaa tactatagcc cacaaaatta tccttcataa atgaaggaga aataaaaggt	12180
ttcccagaca cgaaaatgct gaggtagttt gttactacta gactggacct acaataaatg	12240
ctcaaggggag gtctggaaac tggtagtgaa aggacgacat ttatcatcat gaaaatacat	12300
gaaagtataa aactccctgg taagcaacta aaggaggta tcaaatgtt ccaccagaga	12360
aatctaacta accacaatga caaacaataa gggaaaaaga aaggaacaaa aatatataag	12420
acaacaaata aacaacaata taacaggaag cctcacatat cagtaatcac tttgaatgt	12480
aatgaattac attctccacc taaacgttat gaaatgcctg aatgataaaa ctatatgatc	12540
caaataatatg ctgattacaa gaaacttacc aggcagacat acataggctg aaagtaaaag	12600
aatggtaaaa gatattcctt gcaaatggaa agcaatagtg agcaggagta gctatactt	12660
aattagatca tacagacttt aagtcaaaaa gagtaaaata aaaaagacaa aggatgttat	12720
tatataatga tgagattaac ccagcaatgg gaaataacaa ctctaaatgt atatgcattc	12780
aacactagag aactcagatc cacaaagcaa atattagacc taaagagaga aatagactgc	12840
aatacagtaa tagtggagaa cttcaacact ccacttcag tattagacag ataatctagg	12900
caaaaaatca accagtaaat tttagattt aactagattt tagaccaa at gacctaaca	12960
gacatttaca aaacattcca tccaaaccact gcaaatacgaa atttgtgtca tcagcacatg	13020
aaacaatgtc caagatagac caccatgt taggcaccaa atcatgtctc agcaattttt	13080
taaaagttga aatcatatca catatcttct cagaccactg ttgaataatg cttagaaatca	13140
atgccaagaa taacgttgg aactatacaa atacatgcag attaaacaac atgtcctgg	13200
ttgatcactg ggacaataag gaaattaagc tgaaaatcaa aaaattcttg taacaaataa	13260
agattgaaac ataacatatac aaaaccagtgc gcatacagca aaagcagtgc taagagggaa	13320
gtttatagca ataaatgctt acactgaaaa agtagaaata tttaaaatt agcaacctaa	13380
caatgtgcct gaagaaacta aaaaatcaag aacaaatcaa accccaaaatc agcagaagaa	13440
acacaaaaat aaagatcaga aaagaactaa atcaaataca gactaaaaaa atacaaatga	13500
ttaacaaaac taaaatttgg ttattcaada agataaataa aattgataaa ccgctagata	13560
gactaaacaa ggaaaaagaa tatccaaata aacacaatca aaaacgataa aggagacatt	13620
acaacagatg ccacagaaat aaaaaggatc atcagagact attattaaca actatatgct	13680
gaaaaatgga aaatatacagaa attagataa attccttagaa acttacaacc taccaagctg	13740
ttgcattcagg aagaaataga aaacctgaac atatcagtaa tgattagcaa aattgaatca	13800

gtaataaaaa acatctccc	actctttaa agcttggac caaatagcat cacagccta	13860
ttctaccaat catgcaaaga	agaataccag tcttcttgat gctattacaa taaatcagag	13920
gaaggaattc tctctggctc	attctacatg accagtgtca cttgaaacc aaaacctgac	13980
aaggacacca caaaaagaaa	actacaggcc aataaccatg atgaacacag atgaaaaat	14040
cattaacaaa atactggcaa	acggaatcca acgcacatc aaaaaataa tataccacaa	14100
tccagagggt ttgttatcaag	gatacaga tgactcaatg taaataaatac aataaacatg	14160
ataagcatct tcacagaata	taagacaaat gaatatatga tcatctcaat agatgcagaa	14220
aaaaattttt gataaatttc	aacatctttt catgaaaaaa atctctaaaa cttagcatag	14280
aagaaacata cctcaatata	ataaaggcoa tatgtgacaa actcagagct aatatcatac	14340
agaatggggc aaagttaaa	gactttcctc taagaactgg aacaagacaa ggatgcaa	14400
tctcaccact cctatccaca	tagtactaga agtcctagcc aaaacaatca gacaagcaa	14460
agaaataaaa agtatctaa	ttgagaagag caagtaacat tttcctt tgcgtatgtat	14520
atggtttgt atctggaaaa	tactaaaaac tccagcaaaa acctctt tttgattaat	14580
taattttagta aagttcagg	atacaaataa aaaatacaaa agtcagtagc atttctatgc	14640
cccaataata aaatagctag	gaaagaaatc aagaaagtga tcccatattaa attagctaca	14700
aaaaattaaa atacctggga	ataaatcaag gaagttaaag atctctgcac aaaactacaa	14760
aacactgatg aaagaaatta	aggattaaac aaacaaattt agaaacatcc catgtttatg	14820
gatcaaaaga attaatatca	ttaaaatgac catactccc aaagcaattt ccacattcaa	14880
tgcaatttct accaaattac	caatgtcata tttcatagaa ttagaataat cctaaaattt	14940
gtatgaaatg agaacagagc	ccaaatagcc aaagcaattt tgaacataaa gaacaaatct	15000
ggtcctgact taatcactat	gcaatctatg catgtacaa aattgaacat ggatttatc	15060
aatttgtaca aataaaaaaa	tgtaaaaaaaaa gaacaaagct ggaggctata gtagccaaaa	15120
cagcatggta ttttagaca	aatggaatgg aatagaagc tcagaaataa agccatata	15180
atatatgtg tgtgtgtgt	tgtgtataca cacatacatg tataatataat gtgtacat	15240
aatgtttct acatgttcta	atatttat tccattccat tatacatatt ccatttctgt	15300
atataaggta tatagaattt	gaagactatc tgccattaaa aagaatgaaa tcctgtgatt	15360
tgccagcaaca tgggtgaaac	tggagttcat tatcttaagt gaaataatct aggacaaaa	15420
agataaaatat cacatgttct	cacttatatg tgggagctaa taacttgatt acatgaaggt	15480

ggagaatgga aaggttagta ggaaacagag actggaaagg atgaatggag gtaggaggg 15540  
 aaggtgaaga gaagagagtt aaaaggtgta aacatatagt taaaagaaat aaattcaatg 15600  
 cttgatagca gagtacagtg actacagtta acaaaatgta ttatactcag gtgatgaaca 15660  
 cctaaatact tgatcactat gcaattatac acgtgtaaca aatcactat gcactatata 15720  
 cgtgtaaaat taaatgcgta caaataaaaa taataaaata ctaatccagt atcattcact 15780  
 gacaatgtta actcaggtgg ataggcatta agtcaatact actataagaa ccacttcttg 15840  
 tttatgttaa tgccatatacg aatgaaataa aattcactaa aatccaaaaa attagaaaaa 15900  
 ctatcaaaac tcaataatata taagacaacc caataaaaat gtggtcaaag gatttgaaca 15960  
 tacatgtcac caaaaaatattt attcaaattt ccaataaata catgtaacaa tgttcgacat 16020  
 cgtagtcac cagagaaata caaaataaaa tggtaatgag atactactag ataggcttt 16080  
 acagagactg acaataccaa gtattgacaa ggatatggag caactgaaat tctcattcct 16140  
 tgtggtaaga atgtacaatt atataaccac attgaaaaaa caagtttca gtttcttat 16200  
 tcacccaaaa tatatgtctt ttggaaaaaa tttttccag tctgtgggtt gtcttctcat 16260  
 tctcttgata tatgtctttt caaagaggct gagcttact ttagacagtg gtcataaag 16320  
 tgtgtatatt tgtgtttta taatttatac gcatatattt ctgtgaaaag atactgtatg 16380  
 cattgttcaa catgtacaaa tataagaaag atatagtaaa gaaatatata tttctaaatt 16440  
 tataaatgta tttattggtg ttccacgttt caaactaaat aatctacgtt ggctaattt 16500  
 aggaattaaa ctatagtaga aggttctcat ttattggat gattagaacc agccttttg 16560  
 caggctatta gcgaatcata gcactaggc ttcaactgcta cctccactga cacctctgac 16620  
 acttgaaact tgagggcaga tatctgccca tgctgataga aaacaactga ataatttaat 16680  
 ttgctagata atagaaaaga atcaaattgac tctgccacat tgcttgccag aagattgttt 16740  
 ttctcatttg tgacctcttg cctataaatg atagatagtc cctgtgctgc atgctatagg 16800  
 tgttcgtaag agagtcgtt aatgtgagct ttttatatcc tattttggg tggtaaagg 16860  
 cattctatta gtctgttctt aaactgctaa tgaagacata ccccaaattt ggtactttat 16920  
 gaaagaaaga ggttaattt actcacagtt caacatgact ggggaggccct aaggaaagtt 16980  
 ataatcatgg gggaaaggaa agcacacatg tccttcacat ggtacgagga aggataatga 17040  
 gtaaaagggg gaaaagcccc ttataaaaact atcaaattccc atgagaactc actctcaca 17100  
 gaacacaaattt agagtaactg ccccatgac tcaattactt cccaccaggc ccctcccaca 17160  
 acacatgggg cttatggaa ctacaattca agatgagatt tgggtgggaa cacagccaca 17220

ccatttcatt ccacctctga cccctccaa atctcggtt ctcacaattc aaatacaatc 17280  
 atgcccttcc aacagtcccc ccaaagtctt aacacattc agtattaaca caaaagtcca 17340  
 agtccaaagt ctaatctgag acaaggcaag tccctctgc ctatgagcct gtaaattcga 17400  
 aagcaagttt gctacttcct agataacaata gggcacagt cattgggtaa atacacacat 17460  
 tccaaacggg aggaattgac caaaaccaag gggctacagg cctcatggag gtccaaaatc 17520  
 caatagggcc attgttaaac cttaaagttt caaaatttac tccttgact tcatactca 17580  
 cgtctaggc atgattatgc aagaggtggg ctcccacagc tttggcagc tctgcctctg 17640  
 tggcttgca gggcacagcc ccactccagg ctgctttac aagctagtgt tgagtgcctg 17700  
 cagctttcc aggcacatgg gtgcaagctg taggtggatc taccattctg tggctggag 17760  
 gatggtggcc ttcatctcac agatccacta ggcagttaccc cagtggggac tctgtgtgg 17820  
 ggctctgatc ccacatttcc cttccacact gcccttagcag aggttcacca tgagggctcc 17880  
 acccctgcag caaacttctg cctgaacatc caagcatttcc ttacatcct ctgaaatcta 17940  
 ggcggaggtt tccagacctc aattgttgc ttctctgcaa atgtaggctc aacacccat 18000  
 ggaagctggc aaagcttggg gcttcaccc tctgaagcca tggccttagc tgtaccttgg 18060  
 cccttattag ttaaagctgg agcagctggg ttgcagggca ccaagtcct atggtcata 18120  
 cagcaggggg gccctggacc cagcccacaa aaccaattt ccctcctagg cttctggcc 18180  
 tgcgtatgagt agggttgcca caaaactgtc tgacatgcct tggagacatt ttccctattg 18240  
 tcttattaag atttggctca tagttactta tgcaaatttc tgcagcaggc ttgaatttct 18300  
 cctcagaaaa ttagttttc tttctatgg catcatcagg ttgcaaattt ttaaaacttt 18360  
 tatgctctgc ttcccttta caattaagtt ccaattccaa accatatctt tctggataca 18420  
 taaaactgaa tgcttataac agcaccacaa tcatacctg aacactttgc ttctcagaaa 18480  
 tatcttctac cagataccct aaattatcgc tctcaagttc aaagtaccac agatctctag 18540  
 ggcagggggca aaatgccacc agtctctttg ctaaagcata acaagagtca ccttgctcc 18600  
 agttcccaac aagttcctca tctccatctg agaccacctt agcctggatt tcattgtcca 18660  
 tattcattatc agcatgttgg tcaaagccat tcaacaagtc tcttaggaagt ttcaaacttt 18720  
 cccacatctt cctatcttt tctgaggcct ccaaactgtt ccaacttctg cctgttaccc 18780  
 agttgcaaaag ttactgcccac atttctgggt atcttacag cagtgccccca ctccctggtag 18840  
 caatttacca tatccattta ttctcatgct gataataaaag acatacccaa ggctgggtag 18900

tttataaaaga	aaaaagaggt	ttaattgact	cacagttcag	catggttggc	aaggcctcag	18960
gaaacagaat	catggtgaa	gggaagcaaa	cacatcctcc	ttcacatggt	ggcagggaga	19020
agaatgagca	aaacggggga	aaaacccta	taaaatcatc	agatctcatg	agaactcact	19080
ctcttgagaa	cagcatgagg	gtaaccatgt	ccatgattcc	attacctccc	aacgggttcc	19140
tcccatgaca	cgtgaagatt	atgggaacta	ctacaattca	agaggagatt	tgggtgggga	19200
cacagccaaa	ccatgtcagt	catgatatga	gaaattatca	aattaagatg	taggaaggt	19260
tttaaaaga	tttgagcaac	cacaaatgac	agatatgtgc	tatagtagtg	caaatacca	19320
tttgctctt	attaaaaata	taattgttct	tgataatctg	aattataat	gtcatggata	19380
attatgatgc	attatgctct	cagcagctaa	aacttcaagc	aaaatacaca	cctagagagc	19440
aatcagcctt	aacaataatt	ctataaattt	aattttcttt	atttctgata	attacatttt	19500
agttgacttc	atatgtgatc	taaatacatt	accattattt	tggacttatg	atgtagctct	19560
tgaagtacat	atatgtgta	gctctaaag	tacatataga	agagcagata	aagtatcagt	19620
tcaccatttc	ttttagttt	gtgcttcat	gatgaatatt	ctcatcaatg	tacagattat	19680
ttgcaggagc	cttttaaattc	catgtgtcca	ttttatgaga	cttagcttt	gtctgtatat	19740
aatgtgttta	ttcagtgatgc	atggattaat	ttgagagagc	acaggtatgg	gtatcttac	19800
agcagtgccc	cactcctggc	accaatttac	tgtatttagtt	tattctcatg	ctactaataa	19860
agactatata	tcacaataaa	ctgagaacca	gctggtaaat	gagagaactg	tggccacct	19920
tttcattgtg	gagttctcat	tttccttagc	ttatgctgct	tattcaacac	tatttctgca	19980
taatctaattg	cattcactaa	atgaaggatgc	tgtgttagcc	tccacatgat	attaatacag	20040
cctatttaat	ttatccttct	tttagattaaa	aataaataag	tagtcatgtg	ccacagaatg	20100
acacttcagt	catttggtca	ttgaaggacc	acatctatta	ctgtggtcca	ataagattat	20160
aataacatata	tttcctgta	cattttcatt	gttctgatat	gttttgatac	ataaatgctt	20220
accatcggt	tagagttgcc	tgcagtattc	agtagactaa	catgtgtac	acctaggagc	20280
aacaggctat	accacatacc	tttaggtgtat	agtttaggtta	taccatctag	gtttgtataa	20340
gtacactcta	tgttgttctc	acaatgaaca	aaatcaccta	atgatgcatt	tctaaaaaca	20400
tgtccctgtc	attaatacag	tatgtaccaa	tacagttatg	acaatatgt	atacatgact	20460
atattcagaa	ttttagctat	ttctcttata	tttcaaattgg	attttcttat	gcactgtgt	20520
gcacgggcat	ttcattttag	taaccacagt	ctgggaaagg	agaagtcttt	gaaggatgtt	20580
gagcaagggtt	atgacatggc	cagatgtgaa	tttttgcata	gtgactccat	gttagcagat	20640

aaagttgtat tgggaaagat caaaagcatg aaggccagat aagaggatac tgtatgttat	20700
catggatgga aatgtgaggg atggcaggag agatgctatg attgaatgaa tctcaatatt	20760
cttggtgatc aaagaataat gagactcatc caataagact ctgtaatga ttgaatgttag	20820
ttcctaagct aggaggaaga atgaggaatg attttctggt tcctgactac agcacaagtt	20880
tttgattttt agaacaaga ataaaattgt acatgctta tgattcctgg ttgaattttt	20940
aaggataaaaa aagtcaagctg taatattattt ctttcctgat accatgcagt atttgtatca	21000
gtgatcttat tcattccaca cacattcttc ttgaacctgg acactgctct agacactgat	21060
tctttccaaa tatcagataa ggttattctt acgttagaccc tcagttcata taaatatgtat	21120
tttccccaaa tgtgaaataa gtgacttttc ataagatatt tttaaaaga atgtcttaat	21180
aataaaattgt gaatgttgca tggaaatgta ggtgacttgc attgtgcata ctgtgtttga	21240
ttcactgctc ttgcattgtct tgcctttagc tggatgaca gcagttcagt gagcagtgg	21300
ctcagtgaca cccttgataa catcagcact gatgacactgaa acaccacatc ctctgtcagc	21360
tcttactcca acatcacccgt cccctctagg aagaatactc aggtgagaat taccacctt	21420
ctttttccag tgtttctgcc agcttttcc ccaaattac ttaatattag attaaggat	21480
agcacaagcc cttaatccaa aattattaca gaaactggaa aatgcagaga taataaggac	21540
tccctttgcc actcctgaac cctgaagcat cttcatctt agtcttcctt aaagccacaa	21600
cccttaggag gagcaacaat gtgcactgca gccaattttg aataaacaga agcagctt	21660
atatatatat atatatatat atatatgata tacattacat atttatatat	21720
atgtaatata tggccatata agcctgggtg tatagttatac tatacaaata tattttttt	21780
ttgttaatata atagagtata taaatatcta tttatataat agatattttt atatattaaa	21840
tatctattta tataatagat atttatatat attaaatata taaaaatata taacatataa	21900
tagatata ttttatatat tatataaata tatattata tatttaatata attaatgatg	21960
aattactata ttgtataga taactacacc accaagctat atgggtgtat tatattaata	22020
tataatgtat aattctatat taatataata gtaacatatac aatacttaat ataatataa	22080
ttcaattgtat tacaatctaa ttcaaaaaa gtttgcataat cttacaatat	22140
cgatatgttt gtttggggat ccagcaatta ttttgcataat ctaattttttt atagtttttgc	22200
attaatttttata ttttttttttttata gaaatataatc actggatctg ttgtaatga taaatcaaaa	22260
atgaaaaatg gacattacat cattaaggatcc tagttgtct tactacttct tatgacattt	22320

gatataaaaa atttctacct ttctgttagcg tttaattgggt gttttctgca tgtattttttt	22380
ctgaaattct ctaatatctg caagtggaa ttatgtggct aaaattaata aaatgttaatgt	22440
gaaggtaaat caaaaatagaa tccttgatt tatccagtta tctgaaaagta catttcatttgc	22500
ccttaattca cactttataa attttctac ataaaagttt tctgtaatat ttgtcttttat	22560
agctgaggac agattcagag aaacgctcca ccacagacga gacctggat agtcctgagg	22620
aactgaaaaaa accagaagaa gattttgaca gccatgggg tgctggtggc aagtggaaaga	22680
ctgtgtcctc tggacttcct gaagaccccg agaaggcagg gcagaaaagct tccctgtctg	22740
tttcacagac aggttcctgg agaagaggca tgtctgccc aggagggggc ccatctaggc	22800
agaaaagctgg aacaagtgc ctcaaaaacac ccggtaggct tgctgttgc cagctgttat	22860
gcaaaaagtgc ttactttat tggatccatt caatcttgc ttctctaacc aatagcattt	22920
ctaaaatacc aaattcttat ccatattaaa catggagtca aatagttaaa tagttttct	22980
gtctacgttt cacaaactcg tcatagaagc ccaagtaggg cctatatcta ggcattctct	23040
ggaaaggcctc ctatcaaact aggggtactg gatgccttac cttgccagag ttatccagg	23100
taatggggaa ataagattag gttgctaaag caacagttaa gtttttttgt ttttgttctg	23160
cgttcttaat gaaagttgg aatttttaca cttaaatatgc cactgaattt cactacagac	23220
tctgagagga acaagcaatg acactaatca attggaatgc tggagatttgc aaatattgtc	23280
tgtgttattt acttcatgaa agaagagaat gaaatagtgc ttcaaaaattt tgccatactt	23340
tttttaaaaaa gactctcccc gtatttttaa aataatgcct aattataaat agtgcacact	23400
gaagcactaa ttaacagggt actccaaata taatcatctc acagatatttca aatgaattt	23460
ttttcttagt aattagcttgc atagggttaa gtgttacctt tttaaaaaaga gttgcaaaat	23520
ataagacatt aacaaatagc aaaacatatgc ttatctttt atctcttcca tctctcataa	23580
tgtttcttct gacagccaaa tttttgttagc tatgcactca gtcctctcaa tatatgagat	23640
ttttgatcttca agccaataca tttaggaagg gaaataatataa aagaagcat tcacatttt	23700
cacattgtttt cacgaagtgtt ggtgatatca aactctacag gcacatataat ttgtgttattt	23760
ctccttaattt agggaaaaacc gatgatgcca aagcttctga gaaaggaaaaa gtccttctaa	23820
aaggatcatc tctacaaaga tctccttcag atgcaggaaa aagcagtggaa gatgaaggaa	23880
aaaagcccccc ctcaggcattt ggaagatcga ctgccaccag ctcctttggc tttaagaaac	23940
caagtggagt agggtcatctt gccatgatca ccagcagtgg agcaaccata acaagtggct	24000
ctgcaacactt gggtaaaattt ccaaaaatctg ctgccattgg cgggaagtc aatgcaggaa	24060

gaaaaaccag	tttggacggt	tcacagaatc	aggatgatgt	tgtgctgcat	gttagctcaa	24120
agactaccct	acaatatcgc	agcttgcccc	gcccttcaaa	atccagcacc	agtggcattc	24180
ctggccgagg	aggccacaga	tccagtagcca	gcagtattga	ttccaacgtc	agcagcaagt	24240
ctgctggggc	caccacctcg	aaactgagag	aaccaactaa	aattgggtca	gggcgctcga	24300
gtcctgtcac	cgtcaaccaa	acagacaagg	aaaaggaaaa	agtagcagtc	tcagattcag	24360
aaagtgttcc	tttgcaggt	tcccccaaat	ccagccccac	ctctgccagc	gcctgtggtg	24420
cacaaggctc	caggcagcca	ggatccaagt	atccagatata	tgcctcaccc	acatttcgaa	24480
ggttaaggatg	tataaaatga	tgctggaaaa	atataaagga	taaatatgtt	tttagacacat	24540
acattacata	taaatgtgtg	tatataatata	ttttaaatat	gtataaggta	tataatataat	24600
atatcttaga	attctttaaa	gtacacagtg	agctctatga	agcttatcat	ataaacagct	24660
agcaaaaaaaaa	atagttctca	ttttgagaaa	cagtcaaact	tcaaagtttc	actgtcattg	24720
tgataactagc	aacacaaaaca	tctaagagac	ttaaaagctg	atggttaatac	ctaagtgttag	24780
tgataaggca	aagtaatagc	ttgtaaaatt	tctatagatt	tccattcctc	cttttcacat	24840
ttaaaaattaa	aaccaaataag	gttttcatga	cttttggcat	tcatttccag	tgtcattttc	24900
ttgctggctc	ttaatgagtt	ggtgatcata	aatgtagatg	aagttgtttt	ccttgtaaca	24960
gattccatttgc	gacagatttta	tacagtgtca	tatcttgaca	cattaaagac	aatcaagata	25020
tgacataatt	tgaaactatt	ccagtgttttgc	gtacagtatc	acaactgaag	agtgggctaa	25080
gctttcttaac	tcttcatctg	ctttcttta	catgactctg	gtaaggatca	tgacttggtt	25140
tctgttcctg	gattgttttt	ggtgttaaat	atgtgaagtt	ctgctctaag	atatcactgt	25200
ttttaatac	ccatgtgttttgc	ttaagtggta	ggaaaataaaa	tgcagttaaa	aattggggac	25260
aaatatctaa	acctctctga	gtctgttttgc	tcatctgcaa	aatggtagag	tgtggtttat	25320
agttcattat	gggttcaata	tttttaatgt	ttgtttttat	tctgttgact	aaaccaggaa	25380
ctttgatatc	ttggaaagga	aagattttga	aacatttatt	ttacaataaaa	gcaatttcag	25440
atacctgatt	gtttgaaaaaa	cctaaaggct	ttattcctcc	gtagtaatat	taatgctgca	25500
gaactgtctt	tttaaaaatac	tgattctcat	tggaaagaat	gaattatggc	gtatagggag	25560
agtaaatatt	tctgtttctt	aagtaaaaagc	caatagtgcc	ctcctgtggc	ccattaccta	25620
tgaaacaatt	tctcatattc	gtcataaaat	atttcactgt	aggaaatatg	gatttcatttgc	25680
caactcaatt	agtaatcatt	atgccattac	ttcatatcat	tgtatccat	tatttacata	25740

aatttgattc taccatctgc ttcatttaca aaactaaaat gtttctgaa ctaaactcca	25800
aaatctaaca gcaccagctc tgtttcaaata cactattaaa aatgtatTTT gaatagcact	25860
ggcaactgac ataaaaaccct ttggccctcg ctggggaaaa tacagacaaa ctgacttgtt	25920
gccgacaata tcaatattgt ttccaaccaa ctgctccctg acagtgactc agaccaccag	25980
atactcaaca caactcccta aacttgcttt aagcgttcca tctagattt gaataaactg	26040
tttaaaaatt taaaaataaa aaaaaaaagag aagagctcat ttaagtgttg tctatcgat	26100
gcgtagaagt tgtttcatta taatggttct gtAAATAGGT aacagcaagt atggtcaaAC	26160
tactgacttt gagtgaaagt ctcatgatca cttaaattat gaaaaccagg ggTTTcatg	26220
tttgacttac tttgttcca cccacttccc ctcttccct agtagcagct cagtactgac	26280
ctacccttat atgagagatt ttctgcactt gataaagaag tccaagctta taaaagttca	26340
ttaacataga gacaggaagt gctttgtagt tcagtagatc aaagcacact tggctctgtg	26400
tactgtaacc cgaaatatta aatgtggata tttagcttctt ggaacaactg aagttgttat	26460
ttgttttct tttaggttgt ttggtgccaa ggcaggtggc aaatctgcct ctgcacctaa	26520
tactgagggt gtgaaatctt cctcagtaat gcccagccct agtaccacat tagcgcggca	26580
aggcagtctg gagtcaccgt cgtccggtaC gggcagcatg ggcagtgcTG gtggcctaag	26640
cggcagcagc agccctctct tcaataaacc ctcagactta actacagatg ttataagctt	26700
aagtcaCTG ttggccTCCA gcccagcatc gttcaCTCT ttcacatcaG gtggctcgt	26760
gtggcgtGCC aatatgagca gttcctctgc aggCAGCAAG gataactCCGA gctaccagTC	26820
catgactagC ctccacacga gctctgagTC cattgacCTC cccCTCAGCC atcatggcTC	26880
cttgtctgga ctgaccacag gcactcacga ggtccagAGC ctgctcatga gaacgggtAG	26940
tgtgagatct actctctcaG aaaggTgAGC ttccTGGAG gcattgataa catcttcccc	27000
ctcttccctg cactatgcct aaccccccacc ccattaaatt cccttgattt cactgtgagt	27060
gccccggTgc aaaaagatgt aagactgatg aaaccgggCc ttcatTTGC tctcattacc	27120
aaatttacag aggaatagaa tcattaaagg tagggTgagt ggataatttt gttaatatga	27180
atgcatacat ttatacccaG taggcaatgt gaataaaatt caaggaatgt atttagatAT	27240
tgaatgaggT ctcctgaaga cattttaatg atttggctta agcttcagaa caacactagc	27300
tccttatgt gacttaagca ttggaaaga ccaaattgaa attattctat agttatgctc	27360
agagcaatat gttaaatttG ttccatttGt acttctatga aaaaatagca gatggattgc	27420
tggaaatcc tagttggcT ggttaaaaaaa aaaaaaaaaa tcaattgtca gccatgaatC	27480

attagagaaa attatagtgt cagtgccatt ttcaatagac tgctaaaaaa gtaatcatat	27540
tacaaaagtgt ttctcattgg ctttatatat atatataaac ttaaagtaga ggacatagca	27600
aggcatttct tacctaataat gcttactgtg aagcatccct tttgagcaaa atcactctaa	27660
atttctcct caaagtgate ctctcttgat tatactgtac tgactcttac caccaggaaa	27720
atgtcttaaa accacttctt ttcctgata aatgcaatgc tatttgtctc ttgacataag	27780
taaagcttta aacatggtct tggccacatg tggaaagaaa tactggtcac gtaaaatacc	27840
tgatatatct ttctatgtct tcccctgttt tttttatTTT ttttttattt ttatTTTTA	27900
actctgatAT tcatgtatggc atttattttc tagaccttca gccttactcc cggaatgata	27960
tttttaaaca tcaattaaag cccttagcta gacactctct gcattacgcc agttccccct	28020
taatgttagga tgtcccaatt tgaattccc catTTTCTCT tgactttgta aaatacaaaa	28080
cccaagcaa aacattgtttt ctTCCCTCT ttacttccta ctgccttaac aatgagacag	28140
ggacagccgt gcaaatgggg ctTCCGATG ataaagtaat tttaacacta actaaaatAT	28200
tgggtttcc tatgggtggc tgctaattac aaaatacatt ttccctccta aagaaaaaaaaa	28260
ctgggccaag gcaaacagct cagtgatAGC aaataaaatg taaccatttc cctatggTTT	28320
tgctgttata tgctattata gacagcatac gtaaaagacca gtaagggttc atTTTCCAC	28380
ctaaaatgtc gggcttcctg taaaatctt gattctagtt tcagcacttc taaggtaaat	28440
gggcattttc acatgtcatt tataaaactt ctaatgaatg aattatatta aaatagataa	28500
acaacctata gtttaatga atgtatccta gattgtatgc tcataatgtaa ggattctaaa	28560
tatcaacttg ataaccaaAC caaacatagt gcaaataGGT tatcattttat taaccacaAC	28620
cacCTTCCAC AAAACTGGTC atTTTTAAT tattaagata atctgcaaca agttggccat	28680
ttagccatca gcctatttct tcagcattta gacattaatc ccagattcag aaataaAGTC	28740
aagtaactat ttataaccaA gtaacattca aatcaAAact agatgaaAGA ttggTTAGTT	28800
gcatacgat AACCaaaATG CAGTTTAAT atTTTACTCT AATCTATAATT TTAACTGAAG	28860
tcaataaaat ttTCactatG gaaatacact agaaaatATG CAATTCTTA TTCTTTTAA	28920
gcagatttat ttattgtaca tgTCAGTCT ttgaaatAGG CCAATTCTTAT ttatgttATG	28980
ttatgttatt tattgtttt gaaatggagc ctcactctgt cgctcaggct ggagggcagt	29040
ggTGCATCT CAGCTCATTG CGTCCTCTGC TACCCGAGTT CAAGCAATTG TCATGCCTCA	29100
GCCACCTGAG TAGCTGGGT TATAGGAGCG GACCACCATG CTGGCTAAT TTTGTATT	29160

ttttagaga tgacgttca ccatgttggc caggctggc tcgaactcct gacttcaagc	29220
gatctaccct ccttggcctc ccaaagtgtg gggattacag gtgtgagccg tggcaccagc	29280
ctgaaatagg ccaattttta aaatgggagt attcctacat taaaatggcc aaataaagac	29340
ttttctaaa ataaaactta aactaatttt ggataaatat gtttgcctt tgaggcttaa	29400
taaaaatgcat taatgaatat taagctgtaa aaagtacatg ttaactacat agctatacg	29460
tataatatta atattaatta gtgccttcca gtaaattact agattaaaat aaattttaa	29520
ataagacact gagcttttg tttcttgac aatagaactg caagcaatag caaattgctc	29580
taatccttcc acgtacattt aagaaaagttt atgacctatt gaagagaaaa gtagatctag	29640
tgggtgatac tggcttcatt atggtaatt aattgatcag tagaatgtca gaaatgctaa	29700
gaaaacccaa gaactacacc agagagaaaa tgtgttaatg taaattttaa ggcaagttaa	29760
ttagcgatat ataataaaga tgtatataag ttcatgattt acctgtttgt ctacaatttt	29820
agatgattt ttgatactca tatttaatc ggtagcttt cctatagatt ttaatttttg	29880
tttaaattcc tcttcgttaa attaaataaa ataataaaat acactttta acagttttct	29940
cttctgcagc tgctctaggt cattgggtggc cattgagcca taactagtct atatttgg	30000
tgggtttgt ttcatgtgtc tgactcaact aaattttaa ataatttgc gtaaccaact	30060
ttgcaaattc tgggttggtc tttaaatgtc agatctggca acgctgcctt gacatttctg	30120
cctagaaaact attggctcta ggcagtcagt gtctgtctgc ttcagactgt tgactgaaat	30180
ccccattcgt tttcatgccc tatctggccc ttgctggcat atgagttgc aacctttgg	30240
gatttgcaga aattgtctat gttagaaaat cattaatatc tagattcaaa catatttcta	30300
aataaagctt taaatttata tggtaacttt aaatgtattt attctaattt tttcattaa	30360
attgctcttc atcatataaa tatataattt ttatacaact ggatgagttt ggcagaagaa	30420
taccaacttt tcattttctt tgtggcatta aactttaact tgtacacatg gaaataaata	30480
atccttaaaa tgacttatga ccacataaaat gccttagcac atgtggttca tatttggaga	30540
tttctcatat ttgttcaata taatttattt tgtttggta tccacagtc ttaagaaaac	30600
ttctatagtc aacatataata ctgtaactgg cctctacaca gtataagcaa ttacctaca	30660
tggctattac cgataaagtt aaagttgtat aaagcctttg gatgctttg atttcagtgc	30720
taaataatgg agtacacata gaagaaaaca ttttagcttt ggtttgagtg atcaaatttt	30780
aggtcagcct ttttacattc atgttatatc atccccatta tgcgtatcct gtgtatttaa	30840
ttttgatcat ttgtatgtcct aaaggaagaa agctataatt ctgcaatttt aattaatttt	30900

acactttgct tatccacatg ccagagatta taaaagaaaat ccctaaactt gtcccactta 30960  
gttgttata tccttttcct gtatttttag agaggccatt tcttattttc tctagacata 31020  
gctttcatt ccttcttgc accaattgtg aattccttaa aatagagatg ataaaattta 31080  
tagcctttta aatacctaattt ttatgatttc taaaagatgg tatacgctaa tttcattttaa 31140  
atattcaaattt aaatgataact agaatcaattt aagtttaag caaacattca tatactttc 31200  
ttcacatgtg taaatggaa ataaacatgc cttttattttt aaaaataattt gaagacaaaa 31260  
gataagtattt aaacaacgtt ttataccatc tctgtcaattt ggaagttgtc actctaactt 31320  
agccagagca gatctatctc attttgcattt tgatatcata gcaaaagtct aatcagttgc 31380  
ataggaaagg aaaaactaag atagtattta atcaatagga ttcagaggaa aattatgcta 31440  
atgtgattta atctattttc tagtaatcctt atcactaaac tgcatttgaa ttgtactgca 31500  
tttagaaagga actcaaataat gtgtgacggc aatggacatc ttgtcacctt tagttggcct 31560  
ttttcaatga gttaagcattt atatgtgtt taccaaaaaa ttatttttta tagttcagag 31620  
aaccatttt gtggatgtg taatttggaa gttttgttta cattatgtcc tttaggggaaa 31680  
tctttgtttt aacagcatgc agcttgcacag aaatacacta cccaaaaagg gactaaggta 31740  
tatattcctc tcagcacaat tgctacctct ctgttgcattt gttaactttt gttgtgtct 31800  
ctcttccttc ttgtttgtt tgcaatgttag cacatgcacat tgaggacgaa atcactttta 31860  
attttgcattt gttctctggc ccgaacagttt ggtgagatag ccccttaggtt agagatacta 31920  
gttagagattt ggctgtctc tcaaattttttt taaattccaa tgtgaatatc actattttgc 31980  
agaaaataata ctaaacaacaa acacaaacaa aacaaaaaca aacaaacaaa aaacttgc 32040  
caggcattac tttttgggg gcagcaactt tggtagaatg cagaactcac ttcaacaaat 32100  
taaaaataata ttaactcttc taacttttgc ctatttaggtt catatgcattt caaatattca 32160  
aaacccatgc agtctacaga tgtggcagt taatgttgcattt aggttgaagg atgctacaat 32220  
ctgaatcaaa gaaaacatata tttcatcatc acaggacaaa tgctgttgcattt aaggtgtgat 32280  
ttttatagaa tccttttgcattt aaaatctcaa aattgttttgcattt tgcaggggta 32340  
ctgctatcag atcaatttttgcattt atctgaattt atctaataatc atttaataat ctcacaaaataa 32400  
tttattccatc cataataaaaaa aataaaaaataa aaatttgcattt ttttgcatttgcattt 32460  
tacttttgcattt tgaggaagag atagaatgtt ctactaatgcattt aggtataaca ctgtatgtgat 32520  
atgaaaaagttt ggcttaattttt ggtgctaaaga atttacttgcattt aaaaagaaaaa agaatatact 32580



acacatctct	gaatttagga	cggaggacaa	tgaaacaaga	aatttcactt	tataattac	34380
ctttgtcaaa	ctatcccaga	gcacatcaat	tccatcatga	aagtactt	ttgacattat	34440
ataaaaaatt	agtaatagaa	aacacacaat	ccaaaacctt	atatttcta	aacttcaagt	34500
taatcatcaa	cttctcttag	attttgaag	acctgaaaat	aacataatt	tcaaataaca	34560
gaactcaaac	accatataca	tttgtaatga	ggcacaacag	tcaatttga	gccttgtatt	34620
ttccaggtt	tagctgaata	atcttcactg	ctttcttagc	ttttgccag	tctagttgg	34680
ggactatTTT	gccttactgg	gcctaaacag	agtgtaatat	taaaatatgt	taataagcca	34740
tactgagaat	aagataaaatg	caggTTTcta	actcCTTtagg	gacacaagtg	gggacaacac	34800
attccatgaa	cacaggtgaa	tgaatgcccc	tagTTCTCT	gagttggaca	atttcatgcg	34860
atcattttt	tctctgagge	caaagtctct	ggTTTgatct	tcctagcagc	ttccagaaca	34920
gaaagtgagt	ttactttgtc	tccatattt	ttttctccat	gctcggaat	cccctgctt	34980
cctgatccc	ccacaaaaac	tcccctgagg	atgaagcctt	ggcttccag	gttccaggg	35040
aagcctcgat	tcctggctgg	aggttagttgt	accacactcc	cagagggcta	aatcccataa	35100
acatcatctt	ctgtctttgt	agatcataga	acttttatt	atcatccagg	aagatttctc	35160
tttgaaaca	aggctggaaa	aactttatgt	cagtcctgac	ctgctttta	atgactgcgt	35220
agagggagat	gcccaacttgg	gttgcagaga	ggacagatct	gcagcccc	ttgcagaga	35280
ttgccagaga	aaacatcctg	gcacagccac	aatcacaact	ccattttct	cccgatagct	35340
cctttgctt	gaaactcatt	ggttacttct	ccagtgtttt	caggtctata	ttctccaggt	35400
actccagcac	ctcttccag	ggcttggaca	aaaatacatc	tgtgttggcc	agcatcagt	35460
ccaaggcagc	agcctccaag	ggctcctgca	cccatggacc	acatccacac	agagaagcac	35520
cttgggtcct	caagtgcctc	cctttcttc	ctttctccca	aacctgaagc	ccagacacta	35580
aggggtcaaa	ccctcctggg	ccctgaggg	tccaaggggcc	tcattactt	ttctttttt	35640
cactggaaaa	aaaattctaa	tcatgcacct	acagaagatt	gacattttc	agtaagttgg	35700
actttccagc	tttcagccag	gacaagactc	aaggctatgt	cttttctatt	gcaacccttc	35760
ccactatatt	gagtagggct	tttagcaatt	gaaaacaatt	atttggtca	tggtttcata	35820
taagctaatt	atttcatatc	aaacaccaag	ttttgtttc	ctaacctata	tagtgataag	35880
agaatttacc	tataatgcca	aagaatgtat	agcttttatt	tgcttaaga	tgcagttgat	35940
tttttaaaaa	agcgaaaagc	ctaacactt	aacttcaaaa	aatgaattta	aatgtttgt	36000

gtaggtcata ggaatatgaa aaaatttat acaacatcta aaacacaccc aaatcaccta	36060
aagtgcata agcttgctaa gtacttcatt ttccttatca attctttcat taattgacgt	36120
taatttgcatt agttgactcc ttcttctatt ttccctcacc attattatcc tgattaaatc	36180
cacccatttatttccattttagg aacaaaaaga ctcaccactt aactatgtct gacattggtg	36240
aagtcgttta aacttaattt tcttatctct tgaatggata cataataacct aggttatatt	36300
gtaaaagaatg acggatatacg tttatgtaaa gatggagaag tgtgtaaagac ttgacagatt	36360
ctgccaaatc attatttca ctggaaagca tgccttacac gatcatagag tagcattcat	36420
cagatatgcc ttagctttgt ctacatttaa ttgagtagta attcgcaaca cagtaaccac	36480
aggattttat gtaaaagaca ttcacagatt gtgttttga aagattgtat ttttgaagta	36540
caaaaactatg acattgttat caaggactca tttaccacaa atatcaaata tttgtgcaaa	36600
gataagttta tgctaagatt tgcataaattt aaagttaaca tggcaactga agctaacatg	36660
tccatggtca caatgtgtta aaaaatgaat gttctgttag cacacttggg aatgtatTTT	36720
attacatagt ttccagagtt aaaaacacaat taataaatga aatgtgaatt atacccat	36780
tgacaacaaa gctctctgta gagcttaat gttctaatga attagaaaac cactgatcaa	36840
atacatccct tacatttcat tgctatagaa accaagtctg aaaggtaag ttacccccc	36900
taggatgtgg gttccccccc ttaatctatt gtggttata tcagagatct ctcagctgtg	36960
tcagacaggc catgacttaa gtgacactgc cctcttgatt ctcttcatac tttccaact	37020
acaattcttt ctcctgggt tgctcatctt aacatagctg tatcattttat tgttagacaca	37080
aggtcacttt tgagagtgaa tgggactata ttaataattt ttccaggtat tagtgcaaa	37140
ccctgggcaa tgcaattcat cctccatctc ctccttatat ttatgtgtt accaagttgt	37200
ttttcctgta gactttttt tatcctaaac cttttctta tttcttcatt cacaacttta	37260
attctaattct ctcaaattcaa catttcactt tctgtctgag accttttca gctctaaaac	37320
taaaatccca tcagtgtgct agaccatata gccacctgaa atcaaagttt tttcttaagt	37380
tctttcttc tatttgtctt ataatttcat gtatcatcct tctctctact ctgcacaaaa	37440
atctgtgtaa tcaatagtct tacttggaaac tgtgtcttc atattgtaca tttcaatag	37500
acaggaacct gtgattttat cttcagaata tctccatcat ctgtctctca tttcaggga	37560
cattgtcattt gctgaagctt ttttaactat agacaattgc agcagatttt aaactgtatct	37620
tactctgtcg actcccttattt gtttcaacat ttccacccat tggaggtat aaaagaagat	37680
attcctgtcc gtgtcaacat aatctcatgt acctctccag atcttagaaaa cacgtatggc	37740

ttcaaatcag gcattggag atcttatgc tgtatggttt cagagtggaa aaaatgattg	37800
attcaaaaac ataataattta aagagttttt attgtattta cagttcacct gaacctctgt	37860
tcattggca agaaaatgag tactctaaa atgcaataat aaattaaagt tactttatta	37920
ttaaatttta aatatatata tatatactta ccttaaatat gtcctcttgt tgtcttttag	37980
catcacccat ttttgatttg accattatct tttctgaata atcagtaaga tacaggatta	38040
ttattaatgt tcaaaagttg cagtattcat gtttcttta ttcttctac caattaaaat	38100
gtgttaatat ataaaatttt tagaaatttt actataaaaa atcacaacat atattagaaa	38160
attaagatca ctacaatatg tcataattag tagactactg tgagctactg ccacagtaaa	38220
ctatggttcg tgtgtcggtcc caagcatgct agccctagta gaaaccatc ccattcaaga	38280
aagactaaca aagtatact tacataaattc aaaaagtctt tggatgaaac ttcatttggg	38340
aaaataaccc aatcgctacc cttcaattttt ttatgaatga aaaaatggaa gaataaaggc	38400
ctctaagatc cattcaaagc caggagacac acaagaattt ctaaatagaa gagaaacaga	38460
agaggtcata gttcttgtga gccatctcat aacctggta gactcattgt catgcctcca	38520
tgcataatcaa caatcgctca gattcattttt tcataatgcc acaagggtta catgcaggaa	38580
cattaatgtc aacctgtcac ttctaatatc catctaataat tctctaaattt cgatggatcc	38640
ttttgcataat ggtgattgtt aaacacctttt gcataggaac agtttctatg cttttgtact	38700
caaataatccc tctacccgttga atcctttccc atcttcgtgt tcaacccatca atcttcgtcag	38760
aatgaactcc tgtcttctat tctttcgaa gcatagaatc tcacggtcag aagagaccac	38820
atctgggtca acccttcatc tcttatgtaa aattttatga catctctagc ttcttcatttta	38880
aacccaccaa tgacagaaac tactaaaatc tagaaataac acctttgaaa ttcttcatttta	38940
aagagatcaa ataaaattttt cctgaatctt cacctattgt tcctagttat atatatccag	39000
attctacaaa ataagtcaaa gttagattgc atatgacagc tcttcataattttaaaaacaata	39060
taataaactc actagttat gtctagctgt agatgcaaaa gtagagatg acttgggtt	39120
atttaaaaac ccagtccagc cagacacattt ggatcatgcc tgtaataccca gcagcactca	39180
ggaggctggg gcaagaggat cccttgcgttca ggagttacag gctacagtga gctatgtcg	39240
tggcactgca tactccagcc tggaaagacag agtgagaccc tgtctcacaa taatagtatt	39300
taataatatc ataaaaaccc agtccacattt tatataggat cctgtttcc tcaagttact	39360
acaaataat atataatctt aataaaaggt tagtggctttt gccaagatag tggcttggct	39420

atgcaaatgc aatttaagac aaagttggta gccctttt tcctaataca ttgccatatc	39480
tgtttctctt ctatttgaa attcttgtgt gtctttggc ttcgaatgga tcttatagtc	39540
cttttattct tccattttt agtcataaaa aaactgaagg gtagtgattg ggttatttgc	39600
ccaaaggcaga tgaaaagcaa aactaccact agaagcttt taccaatttgc tgttccattc	39660
aaaaaaattat ctttgtatgt cttacatttgc tcttctactg tatagttttt cttgttctat	39720
tttacatatt aacttttctc cttttcaga catctgccct actggctact cttgaaatca	39780
gagactgtgt catatttttc cttctattca actacaacat ctaaaaggcag atctgtcata	39840
gttattaact taattgaaca ctcttaaata gtttaggtgtt atttccaatg cagaagctat	39900
caaaaagggtt tgaaatgca aactattccc tttaaaatct atcctaattcc tcattaatgt	39960
ttcatcttgc tagagctaag tattatgtat tgaaattgtt gaagtagact tcacttgat	40020
atctctgcaatc tcatttaggt aagaattata caaaggccaaa aagcaaataa aatatcctcc	40080
taaccctata gatacgatata ctAAAatgtat gcacttgcaat ttgtttaa tacttcattt	40140
atttaaacaa gagtaaatttca atactgtgaa ccaagaatag ggtgacttac cccaatcttg	40200
ccacacctaaa cataaacattt ttaagtcttc aatgtcctac agtgtaccta ctggctgttg	40260
tcactaatca gaccgaaatg gtactaatgg tcactgcagg ctgaaggaat atgcttgaaa	40320
gataggcaga tcctctccct ctcccccccc tactttttc gcctttccat cctttcttct	40380
ttttttccaa tagattgtgc actttggaga ttcatattttt cttccctttc cattacattt	40440
taaatatgtt attcttagtc ctatgcttcc ttttactcca atcaataact ggctctatca	40500
gaggggtgtt ctgtgtgttta attcggttaa taccaggatt atcaagcaca gtgccttcca	40560
aatgtgagat acttctctcc ggttacctct gggtttactt ttccctgtttt acattgtttt	40620
gagagccagt acttgttatttta agaagaagtt tagtgcctgt gtcacagaaaa aaatcttagt	40680
aaatttgaa gtgatgtcag aacaactcta agccactgac ggattccaca gggttttgaa	40740
aataactcgtt agttccctttt atatcttaag aggctcctgc ctgctttctc atataaccagt	40800
aacaaacttg ctttcttaa atatgagcat tttagaatatc tttctcaatt tttctgtttt	40860
gcttttatttcaaaatccac aactatatttgc ttccatgtt tagttgtaca tacaatcaac	40920
caaatttttc cttaaatttgc tgactaccag gtgaggactc tttggcaata agcaataaga	40980
aaataaaatttgc ttataaaaaa ttacagactt aagataacttc tttggaaata taacatgttt	41040
gtgacttttgc accatctcat catgatatgc tcatctttaa cagagtagaa aatcatttca	41100
tataattaac ttatgggtgg gctgcagata ccatgtatgt tacattgtgtt ttagttataa	41160

aaatgttat tatacactat ttccttataa tctaactttg ataataatga tggcctaatt	41220
catgaactta catcaattaa gagcttgaag tgactgagag tattgcctg gaagcattta	41280
aagccttct tggaaattt agatgtttt tatttactt tcttttgat tttgctttt	41340
ccattaaagt gattactatt tttaaagaga aaaccgaaaa ctctagaaag accatcttt	41400
cttcataaca ggttagcagaa aacaccatgt tattacattt ctagcaagag cagtagaggt	41460
gacttggg ttttgttac tggcttta gaaattgatg taaggctcc cataaacgtg	41520
ccagaggaaa agagggacgc aatgggatct gttattgaac atttcagagg cagactctta	41580
ccttaaatag ggactcacta tacattcatg tttcataag tattgggatc atgttcttac	41640
tttctatcaa cctgctattt tcattttca agcttaagag taataggctc tgtgtgttt	41700
gttttcagt gagcccaaca aatttgcctc aatttaacct tcccgggccc agcatgatgc	41760
gctcaaacag catcccagcc caagactctt cttcgatct ctatgatgac tccagcttt	41820
gtgggagtgc cacttctctg gaggaaagac ctgcgtccat cagtcattcg ggctcattca	41880
gagacagcat ggaagaaggt aagcggtgag ggggattaaa gatgaagtca ctttattnaa	41940
accctgagag ggaaaccatc gtgtcactca catcacaaag attcctgaag aggaaaataa	42000
actagtgtaa ttatcatttggaaactaga agcttgaaga agttttatc tgtattatct	42060
tctatttctt tatgtattna gaaatatgcc agaatttgc tatattaata cttggctgt	42120
gaagagtttta gactaaatct actttccaa tacagaaata tacatataaa ctattnccc	42180
aggtgcata aatatcagag caaatgtttt gttgacatt ttggtaaag agccataaaag	42240
acacacaaac cagaaacatt atttatgaa aataccacat gttgctgact tttattncca	42300
ggaattccct ctgggtctaa ttttttatta tatcattttta gaattcatat tgtacctact	42360
ttttgcttt ataagtcaactt atttcttcat ccaatggcaa taaaattgtc acctaaccta	42420
ataaaatatct ttatagttat atagttctat gtaaataactc caaataaaatc agctgaaaa	42480
cctcaggaag ctgagttgat gctcaaataat atatattttt gtaaactgta gaagctcaa	42540
tgtcaaattt aacaataatt tgagagactt ttctcttga ttatgaat ttttttagta	42600
tccataaaga aaacttacag catacatatt ataaagcatg tcagctaagg ataaaataaa	42660
actagacata caaattcaaa ctgatttagaa tgaaattatt aaccctaata attatgttt	42720
aaagaaaagt ctccaaatct tgagacatac cagagttaa gtcttcagcc atccatttac	42780
ttgtgtata aacttaggca agtttcttaa cttcttatac cctaagttct gcatctgtaa	42840

cttcttaggt ttgtcacaag gatgaaatat gagaacaaag aataattctg ttccatgatc	42900
ttttcccttc ctaccttctt atttaaagta tcttctgact gaggggttag gcagcaatga	42960
aaattgactc atgttttca ggtcaccact atggattcaa tatactggca ttaaatcagt	43020
agagaatagt tgtcattgcc ttttgcataa ttaaccaa acactcagttc actgtgacag	43080
acagtgaatt atatccaatg actccactga tttttccat gtagatagac aaaatataac	43140
tactctcaaa tgtaaggacc ctgctttctg aaatggttct gttgctct tcacagatag	43200
gcttcttata atactttaa aataatttgc taagcataca gatggcttc tagagtgtgg	43260
cattgacaaa taaagtgatt tttatatact gggaaattct ggccttcaat gtatcaggat	43320
taaataatct gaatttctga aagctagcct aagtggcaa gatggcttt ttgtgctcac	43380
gcattgaata ctgaactatt cttagttctta aatggcgatc tagattcaag acttattgaa	43440
ctagattgaa gggactttat tgatatccta cctaatgctc acactgacag atgaagagac	43500
ttagccacat gttctaaggt cataaacaga aagaatgaga atgagatggt ctaattaatt	43560
gtccacctt cctatggtaatc atcaggtaa cacttagtt tacgaggta ttattagaga	43620
tagaaagaat tttttttaa ataattgact caaatacaa cattttgcac attacataga	43680
gtaatagctt tgcccaagtt agaaaaactgg gggttcttct ttattcctct tttgaccaca	43740
tctatatact cagttttaaa aaggttcttc ctggtatcct tcaattccat ccccatgttt	43800
tcatctacaa gccttagtgca gctattccag ccgtctcctg atcaggtctt aagcacctcc	43860
catatgtcct tgttagtaccc accatattga tctcagtagc aatcacagta ctctattgta	43920
aatatcttt aaattattat cttctctttt agctttggg attttatctt atttattttt	43980
gtagttccag gatctagcaa cagcttgtca catcggtcat actcaactaa tgtttggta	44040
atgcacaatg agcagaaata aacatactac tccatagtaa aaagaggatg aactttctg	44100
caaataattaa tcagcaccat tttatccacc ttttgggttt agtacattgg aagtatagga	44160
gtataaagca gaatgtccaa tgtttacagt gatatttga aatagataaa agccagtgcg	44220
acatttccat tctcaatttc tctgagacat cacctgaaa aaaaaaaagta ttttctctt	44280
cctaaaattaa gtaaaggaac agtaattcca cattataag agtatgatca acgcatcaca	44340
gataatgttg taataacaca ttagataaaa gtgcttattt tcctgaaatt atatggagaa	44400
aaaaatctga aagtggacct ttgttggata caaatgaaat aaataaggta catacattt	44460
ttaagggtcg aaagttatg gcaactttag tttgggttcc catgctattc tatttattat	44520
atgggaattt actgttagct tcaacatgta cgaaacagggc tggtagggct catgcttgta	44580

ggcttctgtc taataacttg gcaactgagg tactttaggg agtatggatg gggctttcc	44640
atgtctcaac gtcctgactg caaaaaaaaatt atagcaggct gtttctcaga atcttatagt	44700
tagttgttat tacttaattt ccctaaccac ccgttcttta cttttctgt aaaggctgga	44760
attttgagt agaccttatt gttttaactc tattttctg tttgtttct ccagttcatg	44820
gctttcatt atcactggtg tccagcaattt cttctttta ctctacagta agtaatggct	44880
gttaagaaaa agcttgcgt tttgccatgc acacagatga tgaaatagat cattttactg	44940
tgaacagatc acattcatct atgacttgca caggagttgt gtagcaaaat aacggcatac	45000
tctaagctgc ccaataccca ataaagtgcc aggtgctcca cctgccccatc tttggtcact	45060
tacatgtgct ttcacttggc ttttgtgcac tcataat caatgagtgg atgtagaatt	45120
cgtttcata aaacctactg aggtatgact tggagtctct gaaaccatgt atgtatctg	45180
ctatactatc attttagtaa tgacgagttg tccatgtttt gttttttgag ccgtgactgt	45240
taattttttt atagtatccc cttctcattt tttatttta agtttattgt tgagaggatt	45300
atcgaagggt aaaagcagta agggtaaagg gtaaaagcat aaaagaacca gagatgtttt	45360
tttttaata tacctttga aagagtgtga ttttttaac ttttattttt attttatttt	45420
attttatttt ttattttttt ttgagtcgag gtcttgctt gtcacccagg ctggagtaca	45480
atgacacaat catagctcac tgcaaccttg aactcctggg ctcaagttat cttctgcct	45540
cagcctgtca agcagctagg actacaggca cgaccacca tgcccagcta attttaaat	45600
tgttttagag acaaggcat tgctatattt accagactga tcaataccca tggcttcaag	45660
caattcctcc tgcttagcc tccccaaatg ctgggattac aggtgtaaagc cagcacactt	45720
agatagaaac tttattttt aagagaaaaa taccagtgtt tcaagttctt ttgcaaacgt	45780
gtgacattat aattcatttt tgacaaggag agttttctg tttggtaaat acaattctat	45840
ctttttaaa aaagtagcct acaggaagtt atattttatg agtgagtctt ttttagagcta	45900
ggtaaacagt gaggtatatt taaaagcagc ctactgaatc tcaatggac ttgagtacta	45960
tgaataagcc ttaatcctgt actgtaaagg tcatgaagag ttcatagcct ctgctgtcac	46020
tgtcaactg agcatcatgg gcagtatttt tttcactcat tattttagg ttcaatgtt	46080
tgtttgaacc ttctctttt agattaatct catatatttta ctgccttaca tagtattca	46140
aaatctgact gttattggca gaagtaatat tttctaatc tctccttca atgattaaaa	46200
ttacccatag cttctagaaa ttaagaaatc acgatttagtt ttttagttaa tgtactttt	46260

gtgcaaatgg ataaagttagt gatatgtgtaa acacacatga aaaaaacaca taaaagaaat	46320
atattaagac ttatgtttcc tcctgttggg ccagcaactgc catttgttgg ggaattgtat	46380
tctgatttaa accattgcca ttatcatcta tgtgtacat caaaagatgt agcatcatta	46440
ttattctaaa tacatacaat aattaatatt tggataaagc taccttcatg aaacctaaaga	46500
aaaactaaat taaaaagaaa gaaagaaaaga aaaatacact tagatagaag aaataaggc	46560
tagtgattgg tagcacaata gagtgactat agttaacaat aatttattgt acatttcaaa	46620
atagctagaa aagaagattt ggaatgttcc taacaggaag aatgatatt cttccctaaat	46680
gaagaatggg atattccact ttcccagatt tgatcggtac acagcatatg tttgtataat	46740
accacatgca ccccataaaat acatacaact attgtgtatc ccaatattaa agatttttt	46800
gaaaaattta ttcctcaaga aaaggatcat gagtttaaga aaaaacagat tactagtcta	46860
ccagtgtcca gtagacctt ctgtgttaat aaaagtgttc tgtatctaca ctatctaata	46920
tagtaactat gaaccatatg ttgccattga ttatttgaag tatatctggc aaagagatga	46980
attgacttt ttatTTTaaat taatttacat tgaaatagcc acatgtgcct agcagctact	47040
agattggata gtgcaagttt atagagaaca caaggggtac attttagat aggagtggga	47100
tgtcaaaatg atgaggataa tttagaaagca tacatgagaa atattgttt aagagtagaa	47160
tatgaaatgg gaacacagat taaaatagag tatgtatata tatacatata tatgtgtata	47220
tatatacata tgtatgtta tatatacata tatatatgtg tgtgtgtata tatatatatt	47280
tataggccaa tatatggagg tagggtatat cctagtgtta agtgagtaaa gaatggatta	47340
ggtgatcgag ccacatgaga aggtgatatt attagaaaat tgaaagttgt atttgagatg	47400
atgaaaatga tatatttcaa ttgaaaagta aactgttagta aaataattca aataaatgaa	47460
tatTTGGGGA actacttaag agaaaaatca taaaacatga ggagtcatc tttccccagt	47520
ccgccccatgat caggcccttag gatttaattt gcaatgagaa aatacctatg aaaatgcttt	47580
ttaaactatc acatgaaaaa gcaattttt attttcatg ctttcttaat aactctcaat	47640
agagatttag ttgatttgca ttttgcctg gttcaatcaa gaaattatcg cgtgacatca	47700
ggcaagttgc caaatttctt tggactatac ctataaaata aaatttggaa atattagcta	47760
gatctaaccctt atttgtctcc ggatgtctgc aaagtgggtt gaaatcacaa gcctaaccctg	47820
atctgcagag gtgttacctt tggcaaaactt atggttttt gttttgtttt gaaatctaag	47880
gccaagcgcg gtggctcatg ccggtaatct caacactttt ggaggctgag gcgggtggat	47940
cacttgaggt caggagttcg agaccaggctt ggccaaacatg gcaaaacccc gtctctacta	48000



tgagatgtaa attattcaag agtgctttt aaaacagttt tcttataaaag gctatttagga	49740
ttctaccact tagccacttt attathtagc cactatatta ctaagttac atattttaa	49800
aggtagtgaa aatataggga agacaaagct caggtaaaaa gagttctgg caaataaaaat	49860
atatcctgat gtttagacta ctttgctta tgtttctga aagaaaagca gtaaaaaacca	49920
gttcaggtag ttttgtgtca attaatctag aactatacca aaagtagaca tagaaaacga	49980
gagattgttt ttcagcttg gatctgctta tggcaataag cagacttgta ctattcaaca	50040
acattatgca ttcttcaact tttcccagaa taagggagct tcccaaatgc aatggtgcac	50100
ataactcatt ttctggcatt ttgcagcccc gcatgaagaa gaaaaacaga gctaggagtt	50160
ttctggaagt caagtaaaaa acaccctgca aattcctatg gcagtcctcc tttccataag	50220
ctgcatagcc aaaaatgttt gccagacact tttatcactg ggtgtttcag tgtttcatt	50280
gtttaagcgt tttgctgact tgtgataatt aaaatttatta ataatcatta aagaaagaaaa	50340
aagtagaagt aaataatgtt aattatctgt ggttatcagt agaggtctgt atgttacccc	50400
agctttattt gacattgttt gtgatcagta aatcacagaa taaaattctg acatctaaac	50460
cttggctaga ggtctctata attttatgga gtctgtttcc tacaatctgt atgaaagata	50520
cttcaatatt ttaagttac atgcacccat ctttttaga gtataatttt ataactattt	50580
ggtttatgtt gcttagtatt tacatcttag agtctttaa ttctgtctt tgcttaaagg	50640
aatattatgg atcaaatgac ctatattta agaataacctt atggtttata tattaagaaa	50700
catttatata aaattctaaa gtaacttgct tgtactattt caattgaata acttaatgta	50760
tttcattcta ttcttctcat agtagataat aaaaagtaca tcatgattat tgtattcatt	50820
tatacttgcg gaattaattt aaaaatgttt ttatagttaa agtctttctt tttattgttt	50880
tacaggctga agaaaaggct cattcagagg taaaaaaaaa tatgcaatat ttatatttt	50940
tctatttttag tttgcattca tgatgaaatt agtcttgta ccactagagg gctctgtat	51000
acaatagcag aactccacag gactgctgaa gtaaggcagc taattgataa atggctttg	51060
atattgcctc taaaaataa aatgaaagga agtttgtata gcaagctgtc cttcacatt	51120
ctagattgag tcttagctca acacctaata agtttctat aatagtaagc actcatthaag	51180
tcattgataa atgaaggct atggctttcc tattttatta cagtcatttt cccactccct	51240
gtaagaccat ctacacagga taatgggtga aacttgggca ccaaggctcc acaacacagg	51300
atactagcat ctcagactat ctgtttgtg tcattatctt gttgcctcta actgccattt	51360
tatgtgtggt gtgtcaccta ttgttctaat cacatatttc acaaatacat atttggttgc	51420

actcgtgagc aaatcaaact gcattcagga aagaatacta ttttaatttc ccttggtaaa	51480
acatttgc tggtcaaaga gagcaggagg actttaatta tgactttatt caaggtgagg	51540
taatggctgt ttgattggtt tacactgagg caatcagaca acagagaaaa aaaatgcctt	51600
aacaacagct tttgcaaaag tattccccc ctttgaagtc ttatttatt agcccttaaa	51660
aataaaaattt gtgctatgtt taaaaatatt tgaaaattat tgattaaacc aatttgcctt	51720
tataatctct gaaccaaaga gtggatatga ttttaaaaaa tcaaagtgg tttatattca	51780
tcacatggac atgacaaagc ttctaacact gatcatagta tagctactga agcatcgaaa	51840
tgctacatct attgcctta gtagtagtta ttcaactccc cttttatcat tgatgctgta	51900
tcatgagttt tggttaaaaa aaacaatttc aatcacttta cagttccctg gattatattt	51960
taaagatact ggaatcatgt aatagagact attaatttg agaaatgctc tttgagttt	52020
gattcattta tgaataaaaat agacgctgta ttttctgaaa tcattcatag tcattatctt	52080
ataaatgtaa agcaaatgtt attttagact ggggtgtatc tgttccggaa aaaaaaaaaa	52140
acaggaacga agtagaatca catttggtga aattatataa gtgtctactg tttccagctt	52200
agagttctct actttgttag agtgttttag ttgaccacca tttatattca aaaaaatcta	52260
atgcccgggg caaaaactag acagttataa aactatgtca agaattctct ttcaaactgaa	52320
gacagcattc caaaagttca actacaacta tagataagat ttgttttga agaaatgaga	52380
agcatcaaaa gtagaatgtt taacatccaa gtaactgaaa tcccttgaga ctagatatat	52440
acttatagaa cctagtgtaa gattgttata aatgttctat ctttattatg cacaacatga	52500
gacttgcaga acaaactgca gaaagtgc tttaaacatg atataatata	52560
tccttaccct tttctgttcc agttttattt gagtgtgaac ttaactaaaa agaaagatac	52620
cttagaatat acattatatt ggtttatcta attagttgca cctatcattt gtttttccc	52680
ctgattttta agatgtggat aagctataaa gcatctctga gctaataata actcactaaa	52740
taaaggctt gataatacag atttggaaag gcttctctgc agtcatttgc actccagccaa	52800
ataacaattt aaatgtgaac tgattaaatg ttgaattaag cccaaatggg agtgcatttgc	52860
ggatattcca tagcctttga gaagtttca aactatgaga aattaaaaatg tacagaggaa	52920
aaaaaaaaacct aagattttct gaaaaagaac atggagtatc ttttactaaa aaagaacaag	52980
aaaaatatgt gtgtatatac agttttata aagaaaaatat ttttctacag ttttattacc	53040
acagtttttc tagaaggaga agaatcaata cagaggtaa actgctcttgc agtcatttgc	53100

catttggaggg atggcaaatg gagcaagtga gcgtactttg atttgttagat tagagttga 53160  
 cacataacac tttgctttg aatgacattt gcttgtaact gtggagtcag tgttcatatc 53220  
 cttaatttc aggagttgct gctgatacaa tggggtaga atgagctaaa tacagcattt 53280  
 gcttcttgg tttgaattct gggttttaag taaaaatcta cttgcctatt ccattgattt 53340  
 tttaattgc attcagcaaa tccataaaact gcggagagag ctggttgcat cacaagaaaa 53400  
 agttgctacc ctcacatctc agcttcagc aaatgtaagt cacttcattt ttaaaatata 53460  
 ttacaacaaa ttttataga ggaaaatgaa atcattttag taacaaactt acaaattttc 53520  
 agtgcctgat acagacttag attaccaact agcaggactc ataaaaagtt aacattttt 53580  
 gcctactcag taataaaatg taaatccaaa ctgatgagag gcagcaatat ggttttttt 53640  
 gcttgggtt tctaataaga ttggaaacaa tagtaacagc catatgggtt acttctttc 53700  
 ttgtttgcta ttttattac tcctcttgc taagattccc tgacaatgta agaggggtt 53760  
 ttagtgggg actttggaag ataaaatatt cctgtcccc gcctccttca tctcaatgta 53820  
 ttgaacaatt tgttaagcat ccagtttaattt ctaaaatatg aaatttagtc taaataggga 53880  
 tagcttagct gcactgtgga tgagatatgg tttgctcaaa aaaccttggc agccttctca 53940  
 tagcaattta aaagggtaca ctttactgg caccagagca gcccaggatg gcagaaatga 54000  
 tgacaatgaa gaccgtcaat taaattaaca tttactgaat atcttccact gtgtcaggga 54060  
 gcactcagag tagatgcaga atgataaagg agaaatgtgg cactgttccc agtctgagg 54120  
 agcaatggtg ttaagaacag cagtgggggg taagggaaatg cctgctattt tgccatatgt 54180  
 ctacacttc tcactcaaca gtccttgct cagttctgct gcatacgctt gggctgctc 54240  
 tgtgcctccc cacccctccc actgctcctc tactgagttt ttctatctcc tagacaaagc 54300  
 atgatatgtc aagagtgagc aggtgcagac ccacagtgtc agacttgaat aagagccatt 54360  
 tttaaatttt ttttaagcta tcattgtgca atataaatttca taagtatgtc tatcatttca 54420  
 ttcacaatgt attcattttca gcactgtatt tgaattgatt ttatggatggaaatgg 54480  
 gaattaattt tggatttttattt ctatattttttaatagatg gtgttaggag attcctgaaa 54540  
 ataatagcag ttttagata attgtttaaag caaatatgaga aaataagggt attattttac 54600  
 ctgtttgtgt ttttaaagag atagtccaga ggcaaccgtt aattttataa tataggctac 54660  
 atgtatagaa gtatgaaata ttgttgcata ggttctgaa tttgtaccca gaggaagtag 54720  
 aataatgtaa atgtcagaac ctcctgggtt gtgttatct gcaataagaa aggctcaatg 54780  
 gcaaaccctta tttatttagat tgcaggata cttgcagatg tcttgaatga ttactcaggg 54840

tttcatttta ttttaatgt cccttggttg agtcatcat ataattcaga tattggaata	54900
ataaaatggct gctagacata gtggaagatg ggctgatact ttccatttga aatgtaatga	54960
tgcttattgt cttcaaaaaga aaaaactaaa atggtatttc acatttttt gttttgttt	55020
ttgtttttt ttctctgaga atctcattct tactcatgat tattggtttc ttgtgtacca	55080
tttcaacatt tttctattat atgctaatgt gtatatatac ttaatacacaca cgtgaaaaag	55140
cttccacaca cacacacaca cacacacaca cacacataca cacacatacg	55200
gaaccaaatt ctaacatagg ggaataatct tcggagtgaa ctctgtgctg ctgtttgaaa	55260
atggagat aatttttagaa aggttcctgc agttggctac ccacctcgac tgctctaatt	55320
atgcttgtca cactatttc actgatgtgt tttcatgact ttagggcatg aattctcagc	55380
tgggtgttaa tatgaccaac aaagggtgaa aacaggttct tgcattttt taagtactct	55440
tttatgtga aaagcacaga tatgcagata atacataact gaacatccag catatctgtg	55500
gctttaaaat atcacgaaga agagcacaat tagggaaaag aaaacatcta tagtgtttcc	55560
ctaggggaac aatcatttaa aaaaaaataa aaataaggaa cacagactag aagcagcagt	55620
gccaaataga taattcatgc tagtcttgc gttaatttaa aaagtgttag tcttggagac	55680
aaacgccccaa attgctctag gttccactca gctgtatgtg ttatcattag tattaacttt	55740
tgcacgctga tgggagactg atatatatcc tgtttatgt tcctttaaac aatttataat	55800
gtaattttaga aaccttctca aatcacatta gatccacaca aaaacctgta catagcagct	55860
ttatTTTTta atagccaaag aaaggaaaca accaaaaata tcccttaata ggccagttaa	55920
taaacaattt ctgatacatc tatatcatgg actactactc agcaatataa agaaatgact	55980
attgatacgt gcatcaactt gggtgatcc caggggtatt atgctgagtg aaaaaagaca	56040
gttatagaag gtcaaatttt gtataattcc atttatataa cattccagaa atggcaaaat	56100
taaagaaaca gagaacagat tagtgattgc taaggctaa ggatgaagga gagagagagg	56160
tagtgtgact ataggaagag ggagatctt agtttgtat tttgaatgag atggccatca	56220
catgaatcca catatgtcaa tctattaatg taaatcaata ttgtattcct ggctttgata	56280
tataatataa ttttataaga tatataatca ttggggaaa ctggatgaag gatacaaggg	56340
acctccctgt actatcttg caacttcttg tgtatataat tataaaatataat ataatgtatt	56400
aaaatgtata aaataatatt ttaagtatca gatactgatc tttactcagt atatgaagtg	56460
ttctatcata acgtaacatg ctttccttt atttggta ttttagttc aaactaaaat	56520

ataaaatcacc taaagatcta cgacagttct tttgaaaaaaa aatcttgctt ttaatttccc	56580
aggagtttca accttaatcc tctcttttagt gtttctttat ttggtagtga tagggactat	56640
caaagcttct taccatcaaa tacatttact gactaaaaat agaaaaataa tttacattgt	56700
aaaaatgtac aaattgaatg acagtcaaaa ggtacaggta atgaagatat gcattaacat	56760
ctactttaa aaaaaagttt attaaaattc tcttttagac taatgcagta tctggaaatt	56820
tatataaaata gatatgtata taaatgacta ttaaacaatt ttaatgtcag ttatattta	56880
aacattttaa taatattgtt ataactatgg gggtaaaattt ttgtatatat ctgaacattt	56940
ttgttcttaa ggaaataatc atttttacat atccaggaat ttgaattact ctcaagtcac	57000
ctattaatta caagtcattt tgaactcatt cattttcttt gtgttgctt tataatgtca	57060
tttagattt catgcatcat aatcagccat caaataattt agttaatact tgattttcc	57120
tcagttgtaa gaagtgcgtgt gtttaaattt cattcagaat gtttcatttc atctgaatta	57180
atatctgtta atgtatgtaa tatacacata ttttaacat gcatgtactt aaattgatta	57240
tagggacttg gtaaaatttac ttatttata gatattttaa atataatcaa ggattttta	57300
aatctacagt tcccatttga aagtaaaagt aagtcttgc ttactagtt gttcacagta	57360
caagtaaact ttctacctt tggtaaatg tgagtgcagc ccccacagtg agaaattgtt	57420
atattagaac tctaatacgatc ataatttata gggatgaatt tcaatgagtt tggttctaag	57480
aaataatctg ttggttttaa caacattttt aagtatcaga tattcatctt tactcagtt	57540
gtgacatgtta ctctcatagc ttacgtgctt ttcccttatt tggggtgttt tttatatttt	57600
aattggtaata tcgcatattt aaacttggca taattacatt tatatggact ctaaacaata	57660
acttgtattt taatttttaa atttgaatg catctatgtc tctgttaaaa tgcatttctt	57720
tcccttgcc caaatgggt atggtaagtc aagagagtct ctagttact cacctctcat	57780
ttgactggca gagtaaagcc cttgttcagt agaatgtgt ttaagccttc cctcccttt	57840
gtaaagttgt tctgaacaga gctgcataaa accacaggta aagtgttaag ctgattctac	57900
tagcatgtcc ttagaaagga gagcggttat attggcaggt cctattgcct ggcgtttctg	57960
atcaataact caccacaaaa cagaaaaacag aagccgcaca aggaaaggca gaactaaata	58020
aatggtaata gcaaacaata agccagatag cctctggcct ctcgcccaca ccttaaggca	58080
gctgggtcag gtggatgct tttgtttgtc ttttaacgta ttttctttac aaatctcagc	58140
cattacataa tttggaaatg gacacaaggc tagttattac taacattttt aaagacatta	58200
ctgaatgaat gtgtaagaaa acaaaaggc cttttgcct ttcagcagat aagtctttta	58260

accaaaaatc tcttgggtat tttgagattg ttttctactt ctttgcttat ttaatatttt	58320
cataaaatcc gctagttact cttgctttt tgcacatctttt ctaagagaaa acaattgggt	58380
catattatta atgagaaaca cttcagtgtt tggacaattt tttgttagtgg aaaagaaaatg	58440
tgaaaactttt tggtgcagaa tcattcttgg ttcaactaac tactaattttt aaaacataaa	58500
gtcttaata tatataaaatg ttatatgggt aaatatataat tacatataat atatgtttt	58560
tatttataca taatatacta tatatttata catgatatac taaatattttt cccatataaa	58620
taataaaaatg ctctaggcat atatgtgtgt gtgtgtatata atgtatataat atatataacct	58680
tcataacata catatataaaa atactatattt atatataactc taggtataca tatatgccta	58740
tatatgcacc tatatattttt tatattacta tataatataat agtataattt actatataata	58800
ctactatata ttactatata atatataatgta tatataatagt atatattata tagtaatata	58860
ttactatata atatataaat atatgtgtgt atatataat atgccttagag tgtttttaat	58920
ttgtcagtgg gctgtctctg taatctatata gaagaaataaa aatgttagacg ttatgtataa	58980
tgtatattca tcttgggttg tggcatcata gtaattctctt ttacatatactt attcagatta	59040
cttttgcacc agcctaatac attgtatgtat tccaaaacca aagagagtat ggattgaaat	59100
gatattccct ttactaatac tcagtcttgtt ctattttattt acctttatag acttcaccta	59160
acacaagtca gggatattttt atcatcatat taatacaattt ttactctgac cttaaaaat	59220
tgcaactgct aaaggaaaaaa tcagaaccaa ataaaactgtc attaacaacc cccctgaaaa	59280
tccatattttt ttaaaagtca ttttatcaag tctctcagac aagatgtgat accctataag	59340
ttaatcagt ttactttcc atttcttctt cattaagggtt ataaagatta tcatttagtag	59400
aaaaattttc ccttatttgc ctcctttcc atttacccta ttgagtgaga aattttagcct	59460
ctcataactt ctaaaagttagc aatgttaatc tgataaaacta aaccaagggtg agataaaat	59520
aagacaatat tttttttctt caacttttaa gttctggcgt acatggcag gatatgcagg	59580
tttggttacat ggtcaacat atgccatagt gatttgctgc acagatcaac tcattgccta	59640
gatattaagc ccaccatcca ttagcttattc ttccctgattc tctccctccc ctaactcccc	59700
ctgacaggcc ctatgtgtgt tggttccccca ccatgtgccc acgtttctc atcggttctac	59760
tcccacttat aagtgagaag aagtgggttt tggtttctc ttccctgtttt agtttgctga	59820
ggataatggc ttccagctcc atccatgtcc ctccaaagga catgacccatca ttccctttttt	59880
tagctgcata gtattccatg gtgtatatgtt accacattttt ctttatccag tttatcattt	59940

gcatttgggt tgatttcatg tcttgctat tgtgactagt gctgcagtga acataatgca	60000
tgcaggtatc tttataatag aattatttat attcccttgg gtatataaccc agtaatggga	60060
ttactgggtc aatttctgct tccagatctt tgaggaatca tcacactgtc ttccacattg	60120
gttgaactaa tttactctcc caccaacagt gtaaaagcat tccttttct ctgaaacctc	60180
tgcagcacct gttatttctt gactttaata atcaccattc tgactgctgt gagatggtat	60240
ctcattgtgg ttttgcgtt accctttttt ttatatgttt gttggctgca tgactgtctt	60300
cttgtaagtg tctattcata tcctgtctat tcatgtctt gcccaacttt taatgggaa	60360
gtttgtttt tacttgcgca tttgttgaag ttcctttagt actctagata ttagaccttt	60420
gtcaaatgga tagattccac aaatgttctc ccattctgca gattgtctgt tcactctgat	60480
gatagtttct tttgctatgc tgaaggcttt taattagatc ctatgtca actttgctt	60540
ttgttgcata tgcgggat tttttgtca taaaatcttt gcccttacct atgtcttgaa	60600
taatattgcc cagattttgt tctagggttt ttatagttt tggatttac ttgtaagtct	60660
ttaatccatc ttgggttaat ttttgtataa ggtataagga agtggtccag ttttaatttt	60720
ctgtatatgg ctgtcagtt ctaccagcac catttattaa ttgttttcc agttccccca	60780
ttgcttgcgtt ttgtcagggt tgtcgaagat cagatggttg taggtgttt tcactaacat	60840
aatcataaca tacatttcat tgaaaacaac acgactcaaa atgttcttta gtaaccagtt	60900
ataagttttt ttgtgcataa ttacaaactg ccattctaata cataaacatt ttgtggttac	60960
ttatagctag aaaatgtgag taatatagtt tatacagcat actctttaca atcccgattt	61020
ctttgtcaaa cttaattca tattaaattt ataaagtata cacaaagggt aaaggagagt	61080
aattttcttc aagtttcaca tttaaggatt catagtagaa tgattaaacc ttacattct	61140
ccactataag gagaattaaa atggaaatat tgagtaaaat cttacatttc atttagtaag	61200
tgctaataaa gggtttctgc cataattttc cttatattaa aagaaaacac acaatttttag	61260
tttaggttt tagtaaccaa tttatggc atagtgaa tatttctaac aggttaact	61320
gaagtgacca tcatggcat atatatataat tttaaattca catatatgaa tactatacag	61380
taaaaaactaa cttatgtac ataccacatg gatgaatctc aaaacccatg taaagcaaaa	61440
gaaaaccaca aaagaatcat gccatttgat tacacttggg tggttttaa aacaggcata	61500
tctaaacata gtgcttaaa gtgtaaagctt ggtaggaaa aactataaag aaaagcaaga	61560
aaataattac cacagaagtt atgttagaggt tatctttggg gaaggaagag ggaataataa	61620
gagagggaca aagaagagct tcttggttct tgaaatgtcc tatttcttga cttggctggt	61680

gaatgcata atgttcacta tgtgataagt cagggggctg tttcatttt gttca	61740
atataatgtgt ggattttcc acagttgaaa aggtaaagtt caggtgtggt ggctcacacc	61800
tataaatccca gccaacacctt tgccgggcca aggtgggaag aattacttga ggcttaggagt	61860
tggagagtaa cccaggcaac agggtgaggc actgtctcta cagaaaatga aaaaaaaaaaa	61920
aaaaaaaaatgtag ctgggcatgt tggtacatgc ctatagttct tgctacttgg gaggctgagg	61980
caagaggatc actttagccc aggagttaa gcctgcagtg aactagggtt gtggcactgc	62040
actccagcct ggggtggcagc aagacactga gtaaaaagaat aaaataaata attaaaagtt	62100
aaaatatagg aaaaaatgag catagcctta tgctaatttt tcagttacta ggtctgatata	62160
catcacattc cttgcttgtc attgaaaatt ttttaaacta tgatacttt ttttagtgg	62220
atttatccaa ttaaatctgc taacaaattt ggtgtataaa tctcaagggt aagggtatgt	62280
ggagagtgaaa tgggtttgtg tgagagagag agagagaaga gggggaggag aaaaagaagg	62340
aagagggaaag gaatggaaaa agataataaa gagttgttct gatagattaa tcttttagtag	62400
atgtattccc tacaaattgt tttctccat attgcagtgt caggtaaaga aaggcatccc	62460
aggatgaatt cagagctagg aacatgcacc tttgtatcat aatgctaattt gaaggaacat	62520
gtacattcta actgttacca ataatggaat atattccgt tattaagtaa taagctttaa	62580
ttctttgtat ttttgtgatc catttgatag taggtgcctc agcatttcca ctctgctata	62640
agtacatgga gatataatttt atttaagtca tcttattcat gtctttcaaa aagaaattca	62700
tttttggcca aggatttcca aattttgccc catatatagg tatagtttat tatagacttc	62760
gtttgcaaaa tattaaatcc ttatatcctt ttagggacac aataaaattt tataagtttg	62820
agataatgtta cttgcagttc tacctcaggc cgtggtgaga gattgaagtg cctcttcatt	62880
ttaacatttt gggttcaagt tggtgcataa gggcatgcaa atggaaactg gccttatttt	62940
gagctttaat aaaatcgta aatacttctt aatcttaaga gttatagtt tgta	63000
tatgtataat tctctaataat ttAAAACAAA acctgaaagc cacaaaagct tactgtgaaa	63060
ttAAATGTGA tggaatatta ttctctaactg gcttacctgt atttccttca ttgaaggaa	63120
tatgaagtag aaaaAGCCCTT ttattgaaaa gagtttgaa agttaagata actctttca	63180
attcaattct ttgttaagtag aaaaagagta aagataatgt ttagctgtca gcagatgtct	63240
gacacttgat ggagcgtatc attacaatag agcagctaac aatatctgca aaggtcatca	63300
tgaaaagtata aaaaatgagga atatttgcctt attgaccatt tcagtgaccc ttTTTGGGC	63360

ttaagtcta aaaatcttgg cagatcagaa ctatatttc ggcattttga gtgtcaaatc	63420
tctacatgat gtcaagtca gaaggagtta ttacttgcaa aataccatct tcttcagaa	63480
gttaaactca cattaaatgc caggagactg aaacactgat tttaagaaga caaagtttag	63540
aaaagatgaa tgaaaatgtg tgttaaagaa gagtcaccag tcagagctaa ctatgatagt	63600
catagtattt aaagagttgg aacacatgaa attaagcatt ttgtaaaatg aaggctttc	63660
atccatccac ataagattct gacattaaa ctatgttct tccattctgt tcacaggctc	63720
accttgttagc agctttgaa aagagcttag ggaatatgac tggccgattg caaagtctaa	63780
ctatgacagc ggaacaaaag gtatgttcag aaattgccac tggagactga aagaagacag	63840
caaattgcat aggattctt aataatacct gaagctcctt aaaaataata ttccaggctg	63900
agtgcagagg ctcatgcctg taatctcacc actttggag accaagggtgg gtggatcact	63960
taaggtcagg agtgcagagc cagcctggcc aacgtggtaa aatcccattt ctactaaaaa	64020
cacacacaaa aaattagctg ggcattggc cgggtacctg taatcccagc tacgcaggag	64080
gctgaggcag gagaatcact tgaacccagg aggcagagga cgcaatgc caagatcaca	64140
ccactgcact ccagcctggg agacagaaca aaaaaaagag taataataat aaaataatat	64200
tcaattctat actaaattaa aacaatgata ataccttct ttccagattt taatttaaag	64260
attttatcag ttactccat attgaaacac acaaaggcaa acaaaatcct tgctggcag	64320
tctattaatt tacttctgga tggactagt aaaagaatac tgaatgttaa gaaagagaaa	64380
cagtcacata agagaatatt ctggggcaa actgttatgc agttgacaag aatcacactt	64440
tgataagaac ttccacaaat acatggtcac taaatccagc tatagggcat ggctgttagc	64500
taagacacac aggaaggatg cctggactc tgccaagtaa gggacttcag gttacagcag	64560
ctatgaaaca aaggccaatc ctgtgtatt ttgaaataac aagaactagt tgccatctag	64620
ggatatcacc tttgaagaaa agtcatttgt tatataaaaa tacttaaat gAACCTAAAG	64680
gattttatgg tatgaaagaa ggtataccaa aaagaaagga acggagaatt tagttcacga	64740
agacaaatgt attaaaaagg tccatactgc atagaaagcc tggcacctt tcctgtgtatc	64800
accagttac ttacttctct gctgttagtc cagtgccctt aacttccttg gataggatc	64860
agagataggt gaaacctata gaattctatg gagtgtgtgt gtgtgtgtgt gtgcgtgcgt	64920
gtgtgtgtgt gtgtgtgtat gaaaactgta aatgtgcata aatgtcagg tgtccagagc	64980
tttcatctaa ttctcaaaga gaccattat atcagaagtt ttgggtatTT tcaagaatgc	65040
gttcctctat ctatccatag gaatggcttc agtttgcctt ttagattctg taagttatgt	65100

gattagctt acaaaagtag tatgtattac caaattttgt cactttacaa aagtttattt	65160
ttaaaacaga atgaatagtt caatgaaatc aaaagagtaa atcgaatatt ctataattg	65220
ccaagtatta ttagcacatt gtattctctc tcatttcctc cgtataaccct gcccgtgaga	65280
gagaatatta tccattcctg gaaaatctgt tctagcacag ctaacaaact cctttgaaa	65340
cataaatttt ccttcttcc ctcctccct ccctccttcc ctcccttcct tccttttcc	65400
ttttcttcc ttccttcctg cctttttct atccttcctt ttccttcct tacaccctt	65460
cttccttcctt ttccccctct gtctccctct ctttctttt tgctgcagct tgtcaactca	65520
ctatgtata taagaaccca gcaaataaaaaa ttagaaggct ttttagagca gctgacggga	65580
aagaataaaaaa acactggccc coagtattct tgaatgagaa ttctggctat gtctgttaaa	65640
agctgggtaa tcttgagcaa gtttatctaa cctttcttga acctcaaatt caccttccta	65700
aaagtgggga tgataatgac tacctttag gatcaccatg aggagtaaat cagatactgt	65760
tatcatgtca catgctaggg gtcacccaaa aatattacct tcctttacat ttctttttt	65820
cccttgaaaa ttataagata acaccaaatt cctcactggg catataccaa gcatattgtt	65880
ggaaatgagt gtttagaattt aagtctcaat atcttaata agtcaaaatt aatagaattt	65940
ttgtcctcca cccaatattt tcttgaactc tgttatatct gtaagtgaat ttctcatag	66000
aaacatacag agaattttct catatacata tagaaaaaaaaa ttagaggtt tgtaatgt	66060
taatgcctat gattaatgcc tgaatattta aaaataattt ctataacata agagatttt	66120
taatgtgtct acataatcct taaaataaca ttgccaaaat tataaaattt tctcagaaga	66180
tatcagaatg tctcatattt tccttacac ttttttaact gaaaataaaa tcacttcctt	66240
ttgaattgca aactgtatac acacaacaat catggtaac tagtttattt atttgagatt	66300
ataacttgcc tattctcaaa gtgatattta aaagcctata aaattatttg caatgtaaaa	66360
tggtataattt caaagacaga atctaattaa aaccagttaga ataatgtata taacaatata	66420
cctcagccata gataattact actgcaaggc actgaaatga attgaatttc aaggaagcta	66480
tggtacaaag ggagattgtt aggtgtgtt tatttcatt ttctgaccag gagagcataa	66540
tttagactga ggagaaaaact ctttggcact aaattcaagg acgaatttt tgccaaggtt	66600
tttaaattgg ggtcatggaa taacaaaaga caaaatcact gttcaaatacg acatttctct	66660
aaaagctaaag ggcataacat ttaatcatat ttcaactaaag gcatttcctc agggagctga	66720
gataaaaaggg tatattgctc tctggtgatt caacaatcct gagaaaaggc ttgtgaagta	66780

tagagcagag	attcttaaac	tcccttcccc	aagtataag	tttcatttgt	ctatatagtc	66840
attcatcaag	tttatattga	atttgtgctc	ttctaatacgac	aaaacagtac	agacaatata	66900
gatatagaat	gatagatata	ggtctatac	tatagacata	cctatctact	agaactctaa	66960
aagcatatta	tacatgtatg	taatattcct	catggagttt	atatttctca	tatatatctc	67020
atatatatgt	atctctttat	catggagttt	atattttagg	aggcacaga	tgataataaa	67080
aatataatta	aaacaggcca	ggtgtggtga	ctcacacgtg	taatcctagc	actttgaaag	67140
gccaaggcag	gtggactccc	tgagatcagg	agttcaagac	cagcctggcc	aacatagtga	67200
aaccccatct	ctactagaaa	caaaaattag	ccaggcctgg	tggtgggcac	ctgtagtc	67260
agctattcag	gaggttgagg	caggagaatc	acttgaacct	gggaggtgga	ggttgcagta	67320
agccgaggc	atgccactgc	actccagcct	ggcaacaga	gcaagactct	gtctcaaaaa	67380
aaaatataata	tatataatata	atataatata	tatataaata	tatatattat	ataatataata	67440
tataaattac	atatttataa	atatgttaatt	tatataatata	atatataatt	aaaacatata	67500
ggatttcagg	tgatgataag	cactactgaa	aaaagtaaag	ctgagaatga	ggatactgag	67560
aagctggttt	ggaagctaaa	acacaaagta	acaaaggcca	aggtggttac	atgttcttga	67620
ttacataactt	taaaaatgga	taaactaaat	taagactcag	attctagtc	ttgggcttca	67680
cagtgtgatt	ttcagcaatc	acatggcatt	aatagcctga	aactacatca	aaattgtcat	67740
ttgatttata	gaccaaaata	actcccttga	atagagaggg	attcactcct	aacactttc	67800
ctatccag	atgccaata	acacggaatc	tcttgccaaa	tttgtgtggc	agaacactgg	67860
tttatataac	ttatagcctg	gtaagaaaga	aaagacatgt	atgaataact	tagaaggcag	67920
aaaattatca	tgctattaga	ctcagtacaa	tgtcatgtgc	attctcaaag	gaaacatctg	67980
cagaggcagg	agaattgctt	gaaccctgga	ggtgaaggtt	gcactgagct	gagatcatgc	68040
cactgcactc	cagcctgggt	gacagagaga	gactgcatct	caaaaaata	aaaattacaa	68100
aaataaaaaaa	taaaaaatag	tgatcaatct	ggcagcattt	tctgaaagtt	aagcagtatt	68160
cccaatagct	gctaaaagaa	gacatgttat	ataatactaa	gtctgttaagt	aggtaaaaat	68220
taagagaatt	gttaatgtgc	ttgctggga	gtgaaattat	ctctaggcat	taccctatac	68280
ctaacctagg	actcagtaca	ctatgatatt	ggcgtagttt	gaccaagaat	tttacctga	68340
tttcagatcg	ttttctcttc	accagcactt	cttcaccagg	attatatgaa	aaaaattaaa	68400
cctgatgccc	tgagggatcc	attatatgtg	ctgaaataac	ttctttctc	accatctaga	68460
atggtaactag	ctatgtacca	ctcttgtcag	aatcaaggaa	attgctactc	aaatcattgt	68520

gcagcttaat tttctcacag aaggccagtt gagaaaggct caacttctag gaatccagca	68580
aactatattt tttataagta acatTTTAC agaactactt ctaaatcctt gtgttcaaAT	68640
ttactaaAGC tatattcaca gctaaatatt tcagaattta aaatttaaaa gactttcaAA	68700
ttagttccCT gttagCTGTCA tgccaaggca attagaacat atgttaaggt atgaggggtt	68760
tttcttgTTA gaaggtcaga gcagggcaga gaagtagccc cttgtatgag tgatgaagCT	68820
cagatattGA ctcctatgct aaccataaAG cctagtagtt tgctcatttg ttacctctCT	68880
gaaacatttt tttgggtgac tacaaaacAG gaattgaaAC cttcaAAATA agggaaATTG	68940
aaacccaATC tttgAAAATA gataatgCTG caactaaaaAA tttagttgAA taagatttt	69000
acattaACTC tccctaattt acgttatgat atttGCCATC tagaagtgtt ttAAAAAAAT	69060
atattgCTGG agtcagatGA tgcattCCATT aatCTTGGG gcatagaATA atgtgaatCT	69120
aaaattttCA aattatttAC actactggTA tttggTCAT gtaatttatt tgAAactAGA	69180
tgcaatAGGG atggccAGGT tatttcAGTA gaacaactAG caagacttCA gatgcattGG	69240
ggagtgggGA aaggaggACC tgTTTAAGGA aactagAGCT gggaaGTGTG agattaACTT	69300
agtGCCATG tgaggacCTA aaaAGCAGAT gtggTggAAA atttaAACAG gcttgCCTAG	69360
aaggTCAGT tagttgatGA cacttgatGA gattgtCCC AGCTTGGGA ttctcaACAA	69420
agtCTTGTt agtgAGAAAT ttggAAAGAG atcaggTATA gttaAGAAAC tgggttggAA	69480
aggCCACCAG gaaaggCGAA tattctgACA caAAatttGA tcattttatt tggaAGCATT	69540
tcaaggCCTGA cctgaacGAA ttgtttAGCC tcagataCAT gcataAAACT gtgAAAAGAG	69600
acattgACTC aatttagCTT cttaACATG agAAacttTC gtggAAAact agaactttAC	69660
aagCTCAGCT ggtgttGGGG gcatcattAT cttgaatAGC tcactggAGG AAAatgAAAT	69720
cttagtttGG ttctcaggTT ttAAAATATC tatcattttt gaaaAGTGTG aagtaACAAA	69780
atatgatCTG attatCTTAT tcctAAAATC cttgcAGAA ttatcccAGC ctcaatCTTC	69840
tctttAGTAT ttaatgAGAA taagAAactG gaaatgACTG aattggAAAGA gtagactttA	69900
aatccatATC ttgatggCAT atacattttt cagtttttt tctaaatGAT taatgaggAT	69960
tctcaAAact tgagtatCTT ctatgtttCC cttcaACATA aagAAattGT atgAAAATAT	70020
tttAAAAtt tctaAtgATT ttatAGTTAG ctatcttGGG aattcattTC taatcatGTA	70080
cctcatccAA actccccACT atggacAAAA ataaaATAAA aatttAGT tgcatctGAA	70140
ggccacATTA caatttCTAT gcattatAGA aacctgAGAA aatgtatCTT aaaaaATAAA	70200

tgtgaacaac taaccataat tatgaagaag aaaaatgaaa actagaaata aactattgaa	70260
aatgtctat gtatcagtta agttttatt taaaattct ttatgtttat ctctataata	70320
ctattggaa agagagaaag gaaaacctga ctttgcctc atccaaagga ggtgattcca	70380
ctgatttagc caaaataaga ctccctggtt ataataaata ataaagttt tgatgtttt	70440
tatatggac cccactcact aggtgatcag acaccctct gaaaaaaaaaaaaatacg	70500
tatgcaataa agttaaagtt ttatgttatt cttcaaggg gagaaacatc tgtttaacac	70560
agaccagaat atttcaacaa agtcatccca atatttatgg agatcataaa tcaagcgaaa	70620
aaatatattc atcaacaact aaacaaacta cattaaatag tctcaaagca catttcact	70680
tttttctga cagaaaaaca gtttcacaa gtgtggagac attttaccat ggctttaac	70740
agtgaggaag gatgtttaaa taaaggaaa aatttatgg aaagctcaga gaaaagagat	70800
gggtgtggct tgagtgacaa ggtgagagca gatctcatta actgaaatga gagagaagga	70860
aggaattttg caaatatgga aagataacta gtgcaagttt gaacagatta tgtcaatcaa	70920
tgtagaattt ggctatcttt ttaatcaaag aagactatgg aatattttat aggtgtttgc	70980
ttatactcaa agtttaaag aaataacagt atgaatttgg ttgaactaat tttttcata	71040
gataggattc tcccaagtt tatagcatat atattctta actagttatt cttccttttta	71100
catatattgt gccacattga gtaacaacta acctgcta at agctattgg tttaaaaga	71160
taattaatat tagaaagtga tcattttct gttcatatt aaacatgata ttctgaaaaaa	71220
gcaacattgc ctgaatgttc tacattttat cttttgaaa acaggttta taagagattt	71280
cttgtgaaaa gctgaacggtt ctgacactga aataagtcag ctaactcaaa gctaagctta	71340
atttttgac actgttggca tgaggtctca ttcccaattt tttcatttaa agccacaggc	71400
aaatgtttta acagattttta atccgttagta caagcattat tgatcttaa tttaaggata	71460
aaaacctgat tttaattaga atttaatatg cattotagta tttacgttgt ataattaata	71520
tttacattcc atgattccac tatgtaccat ttattcttt ttgaataat ttccagtagg	71580
agcagaataa atttcagtg aatattttat ttcttggggg atattttaa atggaaaata	71640
tattaagttt cggtaaaatc tggtgcta ttggcagtgg acagaatata aaaattggag	71700
agactgagtc attatgatga attgggtctg actttgtca tgacactgga aattccccac	71760
aaatattata ttcttctttt ataataaata tagtcgaaat gaattgcagt caagtatttg	71820
aagacccatc tataaattta ggcggttact gttgattttt cattatgaga gattttcca	71880
ctcataagct actaaaagta cataaagaag gtctgggtgt ttgttttaaa tgtgactgtt	71940

ctctatcagg	aaaatgtcag	gtatccgatg	aaaatagata	tatgaggtgc	caggtatcta	72000
ttccaaactt	ggatatca	tcaattagca	tcatctttt	tttttttaa	agtgtctaag	72060
gttagaatag	tcaccagata	ttcccatgt	tgaagcaatt	ttctgcaaag	gccgctgtgg	72120
atgatctttt	taaaatata	attctggag	acattgagta	aagagaaaatt	atttaccaga	72180
gaatgaagaa	ccgaggccc	attcttggc	tttctgcca	agatgctgaa	ggcagtgt	72240
aatgacaaat	acattacaa	ggaattctcc	ctctaagagg	ctgacaaaaga	tctgatttt	72300
aggattata	taccaccaag	aagatacccc	ttgtca	ctgtcaatg	gaaatatgg	72360
ctatactgaa	acaattctca	gttcttttc	tttctatctt	tttttgagtt	attttatctt	72420
ccaaaaatga	gttatttctg	ataaaataat	tcacttaat	aattatgaaa	gttcaaattt	72480
gtgcaaata	tttatttggg	acatctaaa	attactctaa	attcaaaaag	aaaatata	72540
ctttattaa	atttgatctg	taagctgctt	tgtttgaat	ttaactatta	tataaaaatt	72600
gtataataca	tatattttat	ttactttatt	cctgtgtgc	tttggcttgg	tgagactagg	72660
tctccacatt	aggagttta	ctgaatgaaa	aagtatcaga	atgtaacatg	actttgat	72720
ggcatcagaa	tttaataaga	tgacattaa	taggaattag	ggtaagttc	caggtttac	72780
acttaaatac	aaataatcaa	tttgcagge	acaaaatact	tcaaacaaa	tctgaaatca	72840
ttcatttgc	aaaacttcag	gtttgcagtt	gacaataat	acaatacaat	gcaacagtgc	72900
aatagtgata	tctaaatatc	taatgtatc	ataggtata	ttagtaagt	tgttatctg	72960
aatgagtggt	gtgatatcct	gttttactt	gtactggtga	gttctgggt	ccaccttga	73020
aaggaataaa	gactattcat	atctcttta	taagacaata	agaaaaacaa	acaaacaaac	73080
aaacaaaaaa	ccacccctt	tactttagct	gagaaagaag	ttatttaggt	cagcttgaca	73140
agttcagcta	agcatccaa	tcttcagga	ggttgttact	acataaaatc	aaacctttt	73200
aattcaacta	tgagcagg	gattttattt	ttcttcggg	tactaaagct	tccaaactct	73260
gtttattcc	caggaatctg	aacttata	actaagagaa	accattgaaa	tgctgaaggc	73320
tcagaattct	gctgcccagg	cggctattca	gggagcactg	aatggtccag	accatcctcc	73380
caaaggata	tttagaaatc	atttcatttc	caccaat	aataggcatc	tattttattt	73440
attaattaca	gtagaactgc	atttactcag	tgtca	tgtca	acatactagt	73500
tgttataata	gttgtattaa	tacatactag	tagtattaat	acatactacg	ttggattaa	73560
tgtgatcaga	atcctagaat	tttagaacag	tgacttccat	tatcagataa	tttttaaact	73620

gatcttaaga aatttgggtc tatagttgta tacacatctc tctacttgat tcagtggaga	73680
tggagatgga gtgggttgtt aatacatgca tatctgactt caggcaaaac aaaccatta	73740
ataggtatga taatcttagat ctgtatcaa aatgaaata gtcaatatga tgatatagta	73800
agcagtgggc attgggaaca acttttcctg gatggaggct ataaaaaggt acatttcctg	73860
tagataattt tgaaacaata aaaacaacgg gtgaaaggta gctctgttt aaattattcc	73920
tatgcttaag caattctaaa caatgaaagg ggtatttctg ccactgcccc tacccctggg	73980
ttcaccactg aagaaatgct cattattaat atcgtgtcat tttttcctt tacattggtt	74040
ctatTTactc atttcctgac acttttcaat ggccttcagt gagctcagct ctTcccagc	74100
ttaaaaaatc ctgtcctaaa acatgaatgc cttattatct ctctttcat ttccagaaga	74160
attctgagaa aaattttatg aagtcttc aatgtcttcag ccatttttag accactggag	74220
tgttagctcct tttccctcca ctccacccaa acaatgctct ccaggatcag cagaaactta	74280
catgacacta aattcagtaa aacgtttata attcttattt tattagacag acatggaaac	74340
acgatttgat gctgatattc atttcttcctt atgtgaaaca tccggttttt ctaatgttcg	74400
tgacatcata cattcttggt tttcttctg ttcccttgaa atatTTTTC aatatttctt	74460
ttgttaattc actctttgtt atccatttgtaatttggta tattgttgc tattttttttt aatatttctt	74520
tatgattcta tgcatttat ttaaaatata tagaaaatca tctcataactc tagctgtaat	74580
ttttattat gtgctaatacg ctaataactg tcaaattctag gtctccaggc caggctctgt	74640
atatccagct accaagagag aactccacgt ggatatctt ggatgtctgt tttgcatttt	74700
aaacctaact tctccaaatt tgeacttgctt ttctgtctca gacctgctgc tccttcagtg	74760
ctctttgcct cagtagatag caccaccatc cttccattta gccagaaatc taagtattct	74820
tcataactcc tcctctcctc attgaataaa ttaccaagat ccgttgatcc cattccttaa	74880
atatctcttg gatctgttaa cttttctctg attttactct tgccatccat cacctctctc	74940
ctgaaccatg accacaaacc cctaaatagc cttccctttc ttaatcttat cctgtttac	75000
accagtcttc acgctgaagc cagaatagtc attaagaaac acatctacag gtatcccatt	75060
cattgcctt agaatggaaat acagactcct cagcatgaca taatctctct tcaccagctt	75120
catTTattca acaaataattt attcataacc aattaagtgc cagatgtgc acatataagac	75180
ttcttgcattt gttgttgtcat tgcatttcc atatTCAGC tattttttttt tattttttttt	75240
tattcataag ttctttatga attgtgttca ttccattttgg aatattctac cttgttttgc	75300
cagcataaaag acttttcgag acactgcagc agcagtgaac ctaaatatgt ttcccttgacc	75360

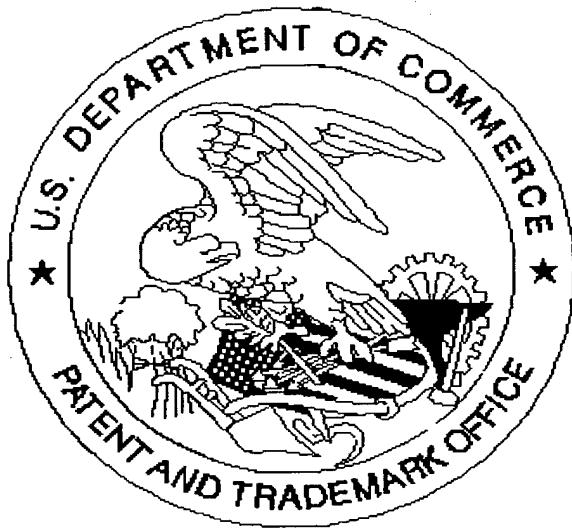
cctacattga atgacaccccc	75420
ctgtgatatg tttctggaag cagcaatact tcccttctta	
aaattacatt atactttggg gccttttattt aaggatatgtc	75480
tttcctgatt tacaatagta	
gagcttgttt ttccaccctt ttgaaaagaca tcaagatgcc	75540
catgatgtg tcttgcatt	
aacaggggtt tatttgaatt tttaaaagaa gaataaagta	75600
atttttaaat gaatttcaat	
ttaaatttta ggaaaacaat tatataaaagt gagatatgct	75660
taaatttgaag gacaaagtag	
ttctgttaggg gctacttctt tcaagacttt agcaacttcc	75720
catgtgggg agtgatttat	
gtgatgcatt gaaaatttact gcattatcaa agcttatctt	75780
agagctataa taaagcagct	
tatgttctaa atcttcatgt cgtaaatagg tccagaaggg	75840
atttaaaaag ccttaatcct	
tactttaaca cagcacaagt cactgaagtg aaacttgctg	75900
aaaggattcc ttttatgtt	
ggcaacaggt agctgaatat atctacagaa attgaaaaat	75960
tggaattctt ttgctcagaa	
atgtgggagg ggtggagctt aaggtaaaaa ataacagtta	76020
atatctaaat tgatcaagaa	
atatgaaaaa ataatttgc	76080
aggtttaaa actaacaaaa accatggta taaaggttt	
aatatatata ggatagttag attgtatttc tgtaatatta	76140
aaactcagca ttaaattttaa	
tgaacacaaa gtgattctt tcacattgac cattgacatt	76200
acatggaaaa aatagtcagt	
tggactaatt atgtgtctt ccatgggta ttaaggtat	76260
tgtatggcat ataaatttat	
actggaaatc acattgaaat tcaactttag aggcccttaa	76320
aatatttctg taatatatat	
ttttaacata tgatctaaa agatatattt ggaatgacac	76380
aacagtttta tagacaggcc	
tgactatcac acaaccacac accaatttgt	76440
aatgtgtttt ctatccctc taaattaatg	
catcacattc attaacaaag tttgataat gactatagtc	76500
tataataaaa tattttgtt	
tacaaacata tttaaacacc tgctattaag tataggcatt	76560
atcagatctt aaaatacaaa	
gattttaaaa attaccctgt ggtcatggag ctcacaatcc	76620
actgcaaaaa taatgtttgt	
gataagaaat ttgaaagttg aaggtaatag aaaatttac	76680
ctttatttt caaaatgtac	
cattgcttcc taagtcacta cttctgtgt aatatggaat	76740
tgtttttcct taagatatac	
caaataatagt tggataacgc atgtattaaa attctgtcag	76800
cactaagttg ttttttagac	
atagtgtat	76860
gcaaacatag ttatattgaa tgaaaaatttta	
gaatcaaattt tattaaacac	
tgtgtactga ttgataccac atgcccatttgc	76920
tttgtatagc aatacaaggt ttgaaattt	
taatggtaaa caaaatagat acggcttttgc	76980
tctccataga acttttagtc tagtgggaga	
gcagaaggta aaggaatgtt tggatcatt ggtgaagctg	77040
aacatgtata cccaaacagt	

tataagttcc aagatggaca ataatgggtg ccatagggaa ggagggtacc aaggaaccta	77100
ctggagggtta catagggaaat attattccaa ggtagtaata tttaagtcaa tatccaagga	77160
ataattgtca atcactttat aagtactgag ggaggagtat ttcaaaagag ctttggcg	77220
gaaaataaat tagttcctt atggaactaa tgtaaggaaa atactaagca aacatgtaat	77280
aagaagaaca cggttgatga gttaagaact gacaagatta ctgaaggatt gtggccata	77340
tttagaagtt ggattttta tctattctta tttaagttag aagttattga aaggtcttaa	77400
gtggggaggt gatgtgaag tttgcctttt aaaaaagatt tttctagcta ttgttatag	77460
aatggttga agatgaataa gtccaatagc tatacttgct gtaaaggta tggttagc	77520
ttgaactggg gcagtggta cacagaggat gggagatgga aaatgacgag tgaacaaaca	77580
catacctgaa aatttaagtt taaaaataga cctctccatt aattcagatt gctgatattc	77640
attcggttag ccattctta ctgaacttta ttagccccca tatactgaat taaatactta	77700
caagcactaa aaaagaaatt gtagggAAC agtaaaatgc atttccttca tttcacaata	77760
ttattaatat tatggctttg ctaatctta ttggtaatg cagtcataat tgaaggtAAC	77820
tgatacttcc aaggactact tttgacctag gattactatc tttttaaaaa ttttagtatta	77880
aagaagtcaa acacaattta ttaattctgg atataataaa aattctgaaa tactttaata	77940
ctttgtgctt ttctatttgt gaaagttaat tattaggaac gagctagcaa atgctacttc	78000
tttttcaaaa agctaattgc caatcacagc aaaaatttaa agcactaaga aataacctaca	78060
catattcttc tattgccat ttatatgact tccataatag ttgattaaag gataccggat	78120
tcctttattt tgaaattaaa acctcctaca tgaaaacctt gatTTAGTT tagaagttgg	78180
taatgttttg gcatgcaaaa ccagttaatg ttctcatcat tacttttaa aacaatgtta	78240
agagatgaat tctaggatt ataaaaaaaaa aaaagctgta tgtgtttttt cctataaaaat	78300
tttcagcat gattgcctca gtagaaaaat taaggactt attgatataat atgtatatga	78360
aggtgaggat acacatatac acacacacat atatatgtag gtaaatacat atattacatg	78420
tctatcaatc catacatact catttattat acgtttgaa agcaaccagt tatagttttg	78480
ttgccatggc tcattttac tattcagtaa atcagtcatt tgaagaggct tgattttatg	78540
gtatttagttt ttggaaact gtcagttta tagtaaattt tgacatctta caacttccac	78600
tgagattttt ttgcttgact aatctgcctt gatgccaata agtataattaa cgaaaaatggaa	78660
ctaaaagcaa atgtgacttg aagcacaatt ttgtaaattt tcttagtgc tcagtaatac	78720
ttaatacttag tgcatttttag gtaggaaaat ttgcattttg ttgtttttta aataactata	78780

aatcttatag ttgcttgtat aaaagaaaaca gatacctta acatgattaa atatcaaatg	78840
ctattctctt caaaatatct taactaaaga agcaactgcct gctcttagaa gttaagcaag	78900
gccataccat atgctgcgta catggctttt aacacaatgg atattagaaa cagcctaagg	78960
ctgagcctgg ctccactatt tttcagctat gtgaccatgt gaaagttaca tttagtaatt	79020
aaactcattt cagtagttt ctttaagaat aaaatttagt actccggggg catatcaagc	79080
atattgtaaa acctagttt attattattt gttattggta ttactattac tattctataa	79140
taagtcatgg gcagggcagta ggggtacatt ggaagaattt cactgtctta aatatgtcct	79200
ctgtttaact cacaaactca gtctacctag gctttcttg gaggatctgc ctttcattgg	79260
ctgtttact ttggccaagt tacttaactt ctttcactt cagtttcctc atctgtgaga	79320
ttatgtgctt acatgacttc aggtttgtt ttggctctaa tatggatgaa ttctatgaaa	79380
tggaaagtta atacatttg ctctagtaac tgtatttcaa gcacaaatat taaaaagcac	79440
aattaattct cattctgagt ttccatttac tctttaaat taatcattca gaataaatca	79500
ttttggaaga gctgcttgat ccaggtattc agtagaaatc actagcatag catttaattt	79560
tagacaaaac tgagaactca ttaaactgcc agggctatgg acttatatga gattctcatt	79620
aaatcttaat gtagataact cagtttaatta aaacaaatat gtttgcattt tattaaactt	79680
ctaaagtcaa aactgcattt aaattatctg tacaaagcct tggatctt tattagagaa	79740
ctgcctctca aaagacctaa aagacttatt tggatcttc gagactcttc atgagccat	79800
gtgatactct ccctcttattt ctagatcttc gcatcagaag acagcattcc tctgaaagtgg	79860
tttcttagt caacagtgc acaagccatt ccagtttgg cagttgtat gatgccgact	79920
ccaagaagaa gaaaaagaaa aactggtaa gttaccatcc ttcatctaa tcagaagctt	79980
attaatgcattt aatgtgttag gccttttct ttggggcttt agtgcatttc agtagtttac	80040
aaagggtccc attcaagcta ctgagacctc aaatgctgca ctcattacca aaattggagt	80100
ggcatgtact gaaaagcata cattttatg ttgggactaa acttgggttt gaatcaccac	80160
tatattctaga ccttttgggg ggcctgaatt ttcttaccaaa taaaaagaca gttaatagca	80220
actatattta tttgtgata tcatttattt acagatgtt aatattttt ctatgtata	80280
actatacaaa ctatgtatgtat taactataga gttatactaa agaaaaataa gataacatct	80340
gtgaataaaat ggcttaaaat aggggtttat tggatcttc gatgtttttt gatgtttttt	80400
aaatgtatgtat gatgtttttt tggatgtttt gatgtttttt gatgtttttt gatgtttttt	80460

gttaataaaag ggaaagaata aacaatgaaa ctctcattcc acctttggaa tcgacagggc 80520  
ttaccgtgtg aatagttca ccctaaaaga aatcaaccac attagtgtct gcttgatgtt 80580  
tttaaccaag agaatatacg agaaatatag aaatgcactt taacagaact gtacctaag 80640  
tttgcttagtg atataattta tgatattgat caatagctaa atagcccagg ggaagatact 80700  
gttactgcga aaaatttaaa aacaatggag tcaatgattt ctttaatac caaaaaaaaaa 80760  
atgtagattt tgagtaaata caactcttga tgaaatccag acataattat cagaggattt 80820  
tactggagtg ctttctacaa ataatgaaag aaatatcttt ttatcttaaa aaatgtttat 80880  
acaggtaata ttttaaaata ctgatcagcc ttcattccct tgatttgtaa ttccacactc 80940  
tttcatgttt ctgcaagggtg aactcttagag gaagtgaggt gaanataaac cgtagacaat 81000  
ttggcatgga tntataaaaaa aaccctacct tggcatgaat gctatccatt ttggcagtag 81060  
gctttatac cttttaaaac agattacctt gtatgtctt tctttgtgtc ttttcatttt 81120  
aatctcaaat tttaaagaga tgtaaaacca ctttctgaat agagctgttag gggataccaa 81180  
ttctggttt gagtagtctg gggttggaaa atttgaatag aaaaatcaca attaatgaag 81240  
tgttaggtga atttgatttc attttgcttt ttaagttgt actgtcagca ggacatgact 81300  
tgattgttagc gctaaagtgg ccatttaaaa caaattgcct tgaagagaga agcattggga 81360  
atggagatc 81369

United States Patent & Trademark Office  
Office of Initial Patent Examination -- Scanning Division



Application deficiencies found during scanning:

Page(s) \_\_\_\_\_ of \_\_\_\_\_ were not present  
for scanning. (Document title)

Page(s) \_\_\_\_\_ of \_\_\_\_\_ were not  
present  
for scanning. (Document title)

*Scanned copy is best available. Some figures are very dark*